RICE UNIVERSITY

Role of Heme in the Folding and Assembly of Globins

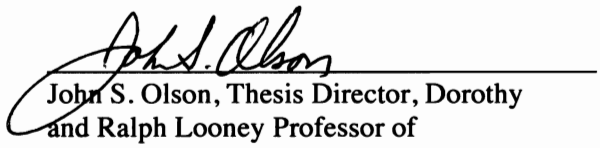
by

David Sean Culbertson

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

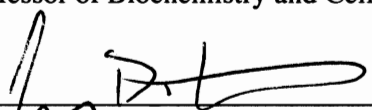
Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

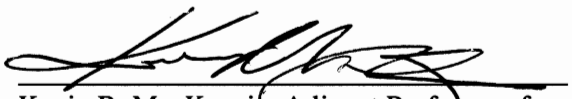


John S. Olson, Thesis Director, Dorothy and Ralph Looney Professor of Biochemistry and Cell Biology

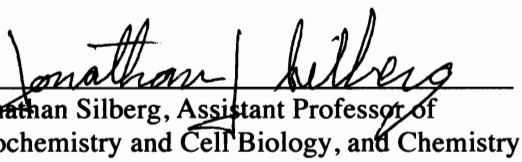
Michael Gustin, Committee Chair, Professor of Biochemistry and Cell Biology



Jeffrey D. Hartgerink, Associate Professor of Chemistry, and Bioengineering



Kevin R. MacKenzie, Adjunct Professor of Biochemistry and Cell Biology



Jonathan Silberg, Assistant Professor of Biochemistry and Cell Biology, and Chemistry

HOUSTON, TEXAS
NOVEMBER 2010

ABSTRACT

Role of Heme in the Folding and Assembly of Globins

by

David Sean Culbertson

Globins constitute a superfamily of proteins that bind a heme cofactor and have diverse physiological roles, ranging from oxygen management to nitric oxide scavenging to gas sensing. Sperm whale myoglobin (Mb) has remained the key system for a wide variety of biophysical investigations of globin function and structure. It possesses the conserved and well-characterized 3-on-3 helical fold, which produces a hydrophobic pocket for heme binding and exogenous ligand coordination. Upon removal of the heme prosthetic group, the protein loses ~40% of its native secondary structure but still serves as a model apo-protein with high helical content for unfolding studies. At least one on-pathway apoMb intermediate is populated kinetically during folding or at equilibrium during acid, urea, or GuHCl-induced denaturation. In contrast, much less is known about the folding and assembly of the holoMb, and how heme confers resistance to denaturation. It is presumed that *in vivo*, oxidation of the heme iron from the ferrous to the ferric state precedes denaturation; however the detailed mechanism for the unfolding of the ferric holoprotein has remained enigmatic. The work presented defines the role of heme quantatively in the folding and assembly of Mb and possibly for the assembly of globins in general. We have demonstrated that ferric holoMb unfolds via a hemin-bound intermediate state, which has the characteristics of a reversible hemichrome species that

is also on the pathway to assembly back to the holoprotein. Similar hemichrome intermediates are observed during unfolding of adult human HbA. The formation of this hemichrome state is very relevant to the biology of erythropoiesis, hemoglobin degradation and the formation of Heinz bodies in circulating red blood cells. We have also tested our methodology using a simpler monomeric hemoglobin from *Cerebratulus lacteus* that possesses low stability and does not populate an intermediate capable of binding heme. All of these results have general implications in biophysics for how cofactor-containing proteins fold, in biology for the understanding of erythropoiesis (*i.e.*, *in vivo* hemoglobin assembly), and in biotechnology for the understanding of how to optimize heterologous expression yields of recombinant holo heme proteins in *E. coli* for either research or commercial purposes.

ACKNOWLEDGEMENTS

The thesis work presented in this dissertation is the fruit of several years of challenging research and I would like to acknowledge all those who have made this accomplishment possible.

This thesis work was supported by U.S. Public Health Service Grants to JSO, GM 35649 and HL 47020, and Grant C-612 from the Robert A. Welch Foundation, Houston, TX. I would also like to acknowledge the support from the predoctoral NIH Traineeship from the Atherosclerosis and Vascular Biology Training Grant T32 HL007812-11GM at Baylor College of Medicine.

I thank the members of the Olson and Wittung-Stafshede laboratories for their support and scientific input, especially Dr. David Apiyo, Dr. Erik Sedlak, Philip Graves, Dr. Michael Perham, Dr. Corey Wilson, Ivan Birukou, Dr. George Blouin, Todd Mollan, Dr. Faiza Hussain, Dr. Agustina Rodriguez-Granillo, Mallory Salter, Eileen Singleton, Dr. Jayashree Soman, Dr. Loren Stagg, and Dr. Neil Varnado.

I thank my thesis advisor Dr. John Olson for being a fantastic mentor and devoting his time to my project. He has guided my research, provided me with help when needed, with always keeping a smile on his face. Dr. Olson has always believed in my ability to solve technical and scientific challenges. I learned a great deal with his guidance and his enthusiasm towards science.

I thank the members of my thesis committee, Dr. Michael Gustin, Dr. Kevin MacKenzie, and Dr. Jonathan Silberg, for those tedious hours of devotion for my project, and Dr. Jeffrey Hartgerink for evaluating my thesis work.

I thank my friends who have been there throughout my life, both in France and in the U.S., and have helped me during difficult times while I was in graduate school.

I thank my parents, Mel and Susan, for always being supportive and proud of my accomplishments.

I want to recognize the relentless support from my wife, Jean, who has always been there to listen and encourage me, and in general has made me a better and stronger person.

Finally, I dedicate this thesis to my brother, Mel, who was a role model to me and would have been proud of my accomplishments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

aquomet	Heme-iron in the ferric (III) state and water as the sixth ligand
CD	Circular Dichroism
CerHb	*Cerebratulus lacteus* Hemoglobin
DNA	Deoxyribose Nucleic Acid
Fluo	Fluorescence
FPLC	Fast Protein Liquid Chromatography
GuHCl	Guanidine hydrochloride
GuSCN	Guanidine thiocyanate
Hb	Hemoglobin
HbA	Hemoglobin A
Heme	Iron-protoporphyrin IX regardless of the ligand and oxidation state
Hemin	Ferric Heme
Mb	Myoglobin
Met	Globin in the ferric Fe(III) state regardless of ligand bound
NMR	Nuclear Magnetic Resonance
Oxy	Heme-iron in the ferrous (II) state and dioxygen as the sixth ligand
PDB	Protein Data Bank (RCSB)
RNA	Ribose Nucleic Acid
sw	sperm whale (*Physeter catodon*)
WT	wild-type

PREFACE

This thesis work was started under the supervision of Dr. Pernilla Wittung-Stafshede who had some experience with myoglobin and electron-transfer triggered folding of redox-proteins during her postdoctoral work under Dr. Harry B. Gray.

The main question was to answer whether holomyoglobin unfolds by a two or three-state mechanism and whether the intermediate state of apomyoglobin can specifically bind heme. After numerous attempts to resolve the key questions using isothermal titration calorimetry, stopped-flow induced refolding, and simple analyses of equilibrium unfolding experiments, it became clear that a more systematic approach using both mutant and wild-type Mbs was needed.

Dr. Pernilla Wittung-Stafshede left Rice University to return to Sweden and Dr. John S. Olson who had been the chair of my thesis committee and is an expert in the globin field generously accepted me into his laboratory. We soon discovered that the apo and holo forms of myoglobin could be compared with a holistic mechanistic approach and that the affinities of the fully folded apo state (N) and partially unfolded apo state (I) for hemin could be obtained quantatively. This approach has proven to be a successful and powerful tool in linking stability of holo protein with stability of the apo protein and cofactor affinity.

Chapter 1

Introduction and background

This chapter is adapted from:

Culbertson, D.S., and Olson, J.S. (2010). Folding and Assembly of Myoglobins and Hemoglobins. In Protein Folding and Metal Ions: Mechanisms, Biology, and Disease, P. Wittung-Stafshede, and C.M. Gomes, eds. (Taylor and Francis, Inc.). pp. 97-122.

Culbertson, D.S., and Olson, J.S. (2010). Role of heme in the unfolding and assembly of myoglobin. Biochemistry *49*, 6052-6063.

1.1. Protein folding and its importance

Proteins are key to any form of life and are the workhorses of cells. They participate in every aspect of what makes life possible, including the structural scaffold of the cell, signaling pathways, cell adhesion, immune responses, enzymatic catalysis of biochemical reactions, and as enzymes are involved in almost every part of cell metabolism. Proteins are made of long chains of amino acids that fold up into a unique configuration that is responsible for their biological function. The information contained in DNA is transcribed into RNA, which is later translated by ribosomes using a pool of 20 different amino acids that differ by size, charge, hydrophobicity, geometry, and polarity. It is the unique amino acid sequence and length that provide their unique function, which can vary from transporting ions through a channel in a membrane to suppressing tumors.

In 1969, Cyrus Levinthal postulated the paradox that if proteins were to sequentially sample every possible conformation to attain the final (native) structure, the

folding process of proteins, even of small sizes, would require an astronomical amount of time, regardless whether each conformation would be sampled every nanosecond (*e.g.*, ~10^{128} years for a 100 amino acid chain) (Zwanzig et al., 1992). The fact that proteins fold spontaneously (*i.e.*, microseconds to seconds) led to the conclusion that proteins fold in a directed process, dictated by energetic constraints, which later resulted in the theoretical representation of the folding process using a funnel-like energy landscape. In 1972, Christian Anfinsen received the Nobel Prize in Chemistry for demonstrating, using ribonuclease, that all the information required for proper folding to reach an energetic and kinetically accessible minimum is contained within the protein primary sequence (Anfinsen, 1972, 1973).

The relationship between structure and pathogenic activity of proteins is a subject of great interest. A growing number of diseases appear to be the result of misfolded proteins (Iannuzzi et al., 2007; Kelly, 1998; Sigurdsson et al., 2002; Westermark et al., 2002). Diseases such as Huntington's, Tay-Sachs, Alzheimer's, Parkinson's, Creutzfeldt-Jacob as well as cystic fibrosis, phenylketonuria, Marfan syndrome, scurvy, hypercholesterolaemia, and sickle cell anemia have been linked to specific proteins misfolding in the ER or cytosol (Dobson, 2002). As a result, scientists are trying to understand how proteins fold to find cures for both genetic and infectious diseases by rescuing a protein's function or alleviating pathologies triggered by protein misfolding.

Proper folding is required for proper activity; yet, more than 30% of proteins in cells require a cofactor to function (Gray, 2003). The addition of a cofactor to the folding process adds another level of complexity to the question of how proteins assemble to become active. Cofactors may bind early on, during, or after folding, and a quantitative

understanding of the role of the cofactor in assembly of their respective proteins is required. Remarkably, the folding and assembly of globins, which contain the heme cofactor, remained poorly understood despite the large amount of literature on the biophysical properties of hemoglobins and myoglobins.

1.2. Globins in biology

Human hemoglobin (Hb) was first identified in 1840 by Hünefeld (1840) as the main pigment in blood, and its primordial role in oxygen transport was later demonstrated by the works of Stokes and Hoppe-Seyler (Severinghaus and Astrup, 1986). Myoglobin (Mb) was later discovered in striated muscle cells and was assigned a role for storing O_2, captured during blood flow and released during muscle contractions (Wittenberg, 1970; Wittenberg and Wittenberg, 2003). Globins have now been discovered in almost all known organisms from all kingdoms of life (Fig. 1.1) (Burmester et al., 2004; Kundu et al., 2003; Vinogradov et al., 2007). Examination of the human genome over the past decade led to the discovery of two additional human globins, neuroglobin and cytoglobin (Fig. 1.3).

The diverse functions of globins

Globins have diverse functions ranging from gas sensing to NO scavenging to O_2 management (Fig. 1.1). The various functions of mammalian Mb and Hb are discussed later in depth. It has been speculated that the animal neuroglobins and cytoglobins provide protection against apoptosis, and facilitate regeneration in neuronal cells and myocytes during hypoxic or ischemic conditions (Brittain et al., 2010; Burmester et al.,

2007; Fordel et al., 2004; Fordel et al., 2007; Hankeln et al., 2005; Schmidt et al., 2004; Singh et al., 2009).

There are also a wide variety of invertebrate animal hemoglobins, which range from small mini-globin monomers (*e.g.*, 109 amino acid globin in *Cerebratulus lacteus* (Pesce et al., 2002; Salter et al., 2008)) to high molecular weight globin aggregates (*e.g.*, the erythrocruorins found in extracellular spaces of polychete and annelid worms (Strand et al., 2004; Vinogradov et al., 1977)). Most of these globins are involved in O_2 scavenging or storage and transport.

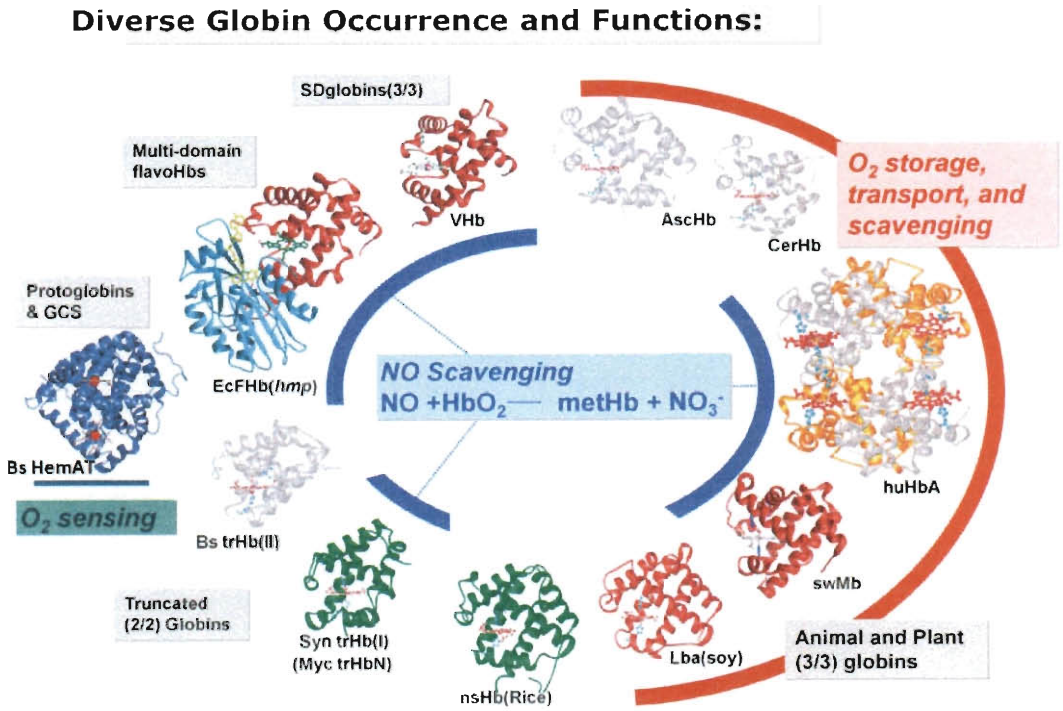


Figure 1.1. The globin family in biology. The functions are depicted when known in red, green, and blue. The following structures were built in ribbons with PyMol from the PDB entries: 1hho for human HbA (HuHBA), 1mgm for sperm whale Mb, 1bin for soybean Leghemoglobin (Lba(soy)), 1d8u for non-symbitiotic rice Hb, 2hz1 for *Synechocystis* Hb (Syn trHb(I)), 1ux8 for *Bacillus anthracis* trHb (Bs trHb(II)), 1or4 for *Bacillus subtilis* HemAT, 1gvh for *E. coli* flavoHb (EcFHb(hmp)), 1vhb for *Vitreoscilla stercoraria* VHb, 1ash for *Ascaris* Hb (AscHb), 1kr7 for *Cerebratulus lacteus* hemoglobin (CerHb).

Several classes of single domain globins are found ubiquitously in plants, the most common, being are non-symbiotic, hexacoordinate hemoglobins, which were first discovered in rice and *Arabidopsis thaliana* (Arredondo-Peter et al., 1997; Kundu et al., 2003; Trevaskis et al., 1997). The exact functions of the hexacoordinate plant Hbs remain unclear but these proteins are thought to be a part of an NO/nitrate cycle that facilitates anaerobic glycolysis in response to hypoxic stress (Dordas, 2009; Dordas et al., 2003). Leghemoglobins evolved from the class 1 non-symbiotic hemoglobins to have a Mb-like function in the root nodules of leguminous plants, where they serve to lower the free concentration of O_2 to ≤ 0.1 μM but still facilitate its diffusion to the nitrogen fixing bacteria (Appleby, 1984; Hoy et al., 2007).

The globin superfamily also includes multidomain bacterial and fungal flavohemoglobins which are fusion proteins consisting of an N-terminal globin domain and a C-terminal flavoprotein ferredoxin:$NADP^+$ reductase domain. Flavohemoglobins are presumed to catalyze $NADH/O_2$ driven dioxygenation as part of a defense mechanism against host-induced nitrosative stress (Gardner et al., 2000; Ullmann et al., 2004).

Globins from prokaryotes such as *Aquifex aeolicus* and *Bacillus anthracis* are presumed to function as O_2-scavengers, possessing very high oxygen affinities; however, these ideas are speculative (Giangiacomo et al., 2005; Miranda et al., 2005). On the other hand, HemAT found in *Bacillus subtilis* has been shown to be a heme-based aerotactic transducer and functions as a gas sensor (Hou et al., 2000; Zhang et al., 2005).

Structure of globins

The basic globin fold consists of several α-helices that are arranged in a pseudo spherical shape and form a pocket for the active site, which binds a heme group. Globins

can be categorized into four groups: (1) protoglobins with more than 8 α helices, (2) truncated 2-on-2 helical Hbs, (3) single domain 3-on-3 helical globins (often with two other short helices), and (4) multiple domain 3/3 helical globins which contain another domain involved in redox reactions or other activities, including protein kinases and DNA binding. Mammalian Mbs, neuroglobins, cytoglobins, and the Hb subunits (α, β, δ, and γ) are classical single domain 3/3 globins (Figs. 1.1-3) with similar folds containing 8 α helices, labeled A to H, with the exception that the α subunit of Hb has only 7 helices and lacks the D helix (Dickerson and Geis, 1983). The characteristic fold is crucial for the formation of a hydrophobic pocket for heme binding (Fig. 1.2-3). Helices C and D are short, variable in length, and located between the longer B and E helices, forming one side of the heme pocket. Sperm whale (sw) Mb has all 8 α helices and serves as the model for the structure and function of all single domain globins.

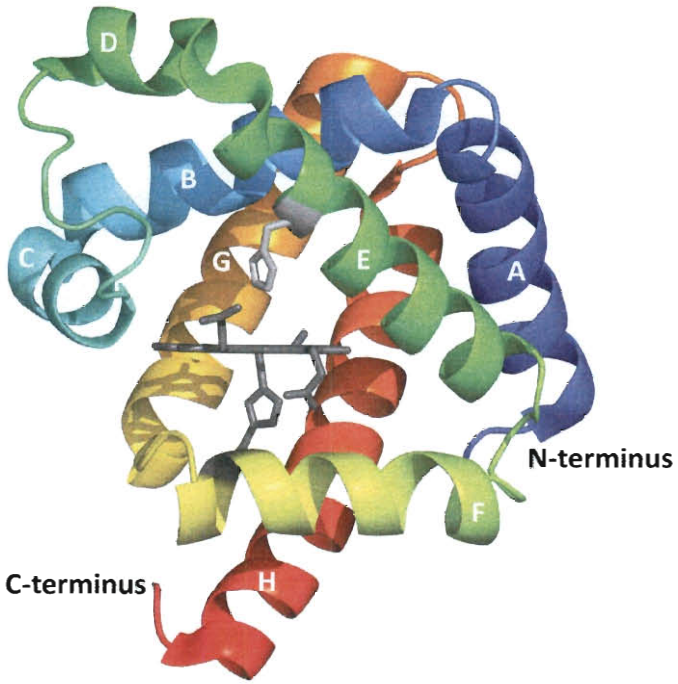


Figure 1.2. Basic helical structure of single domain globins with 3/3 helix fold as found in mammalian Mb and Hb subunits. The proximal and distal histidines coordinate heme (dark sticks), and helices A to H are labeled as well as both termini. Actual figure depicts the crystal structure of native sw metMb. The image was created using PyMol and PDB entry: 1jp6.

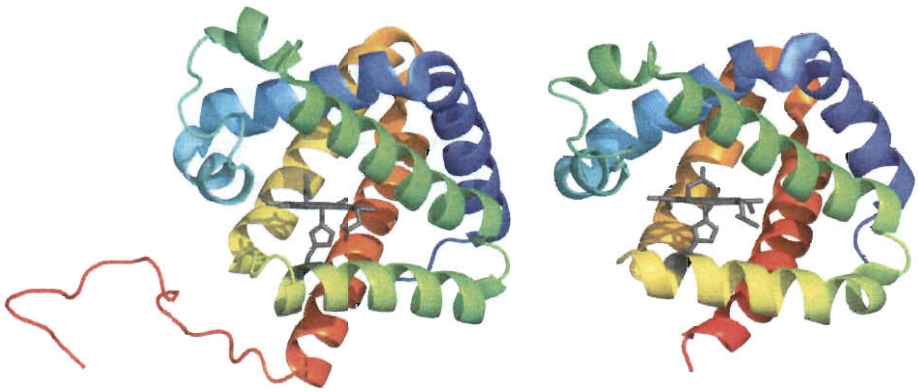


Figure 1.3. Crystal structures of human cytoglobin and neuroglobin. Cytoglobin (left, PDB entry: 2dc3) and neuroglobin (right, PDB entry: 1oj6) are shown in ribbons. The proximal histidines and heme cofactors are shown in gray sticks. Cytoglobin has a C-terminus extension. Images created using PyMol.

Heme coordination in globins

The equatorial *d*-orbitals of the heme-iron coordinate the four tetrapyrrole nitrogens of the pyrroles rings in the equatorial plane, and the axial positions are coordinated to the imidazole side chain of the proximal histidine on the F helix position 8 (F8) and an external 6th ligand when present in the distal portion of the heme pocket (Figs. 1.2 and 1.4). In erythrocytes and myocytes, the heme iron is > 95% reduced (Fe^{II}, ferrous) and is either coordinated to O_2 or pentacoordinate without a sixth ligand (deoxy). *In vitro*, ferrous Mbs and Hbs can bind O_2, CO, NO, or alkyl isocyanides whereas the ferric or met forms (Fe^{III}) can bind H_2O, CN^-, N_3^-, F^-, SCN^-, NO_2^-, or NO (Antonini and Brunori, 1971). The heme-iron may also adopt an oxoferryl Fe^{IV} state after oxidation by hydrogen peroxide, providing peroxidatic activity to globins (Reeder and Wilson, 2005). Binding of ligands is modulated by four factors: the redox state of the heme-iron, steric access to the iron atom on the distal side of the heme group, ease of in-plane movement of the Fe-proximal imidazole complex, and preferential electrostatic interactions with bound O_2 and NO (Olson and Phillips, 1997).

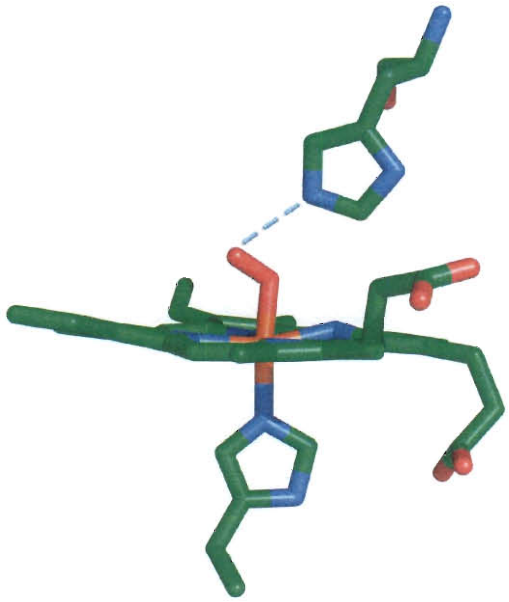


Figure 1.4. Stick representation of heme coordination in sperm whale oxymyoglobin. The heme is coordinated by the proximal histidine His93 (lower) and the distal histidine His64 (upper). The dioxygen is shown in red sticks with the hydrogen bond to the nitrogen of the imidazole ring of H64 shown as a blue dashed line. Image created using PyMol and PBD entry: 1mbo.

Factors governing the stability of globins

The presence of heme in globins confers enhanced resistance to denaturation induced by chemical and thermal perturbations (Hargrove and Olson, 1996; Landfried et al., 2007). Oxidation of Mbs and Hbs to the ferric form disables their capacity to bind O_2 until reconversion to the ferrous forms by intracellular reductases (Xu et al., 1992). The rate of heme loss from ferrous Mb is at least 50 to 100 times slower than that from the ferric form, and holo-globin unfolding does not occur readily at pH values ≥ 7.0 until the heme group is lost (Hargrove et al., 1996b). Thus, under physiological conditions, auto-oxidation almost certainly precedes denaturation.

Some globins have evolved to be able to retain functionality under demanding conditions and need to be resistant to oxidation, heme loss, and denaturation. Sperm whales (*Physeter catodon*), as with other marine mammals that undergo extended periods of apnea when deep diving, possess higher skeletal muscle Mb concentration to increase

O_2 delivery capacity (Kooyman and Ponganis, 1998). The apo-forms of these Mbs display overall folding constants 20 to 100 times greater than their terrestrial mammalian orthologs, making them much easier to study in terms of identifying stable intermediates (Scott et al., 2000).

Oligomerization enhances the stability of proteins both by the strength of the new subunit interfaces and by the multiplication factor for the individual stabilities. For example, the folding constant for an individual subunit is taken to the fourth power when calculating the overall stability of a homo-tetramer as shown in Equation 1.1 below:

$$K_{\text{folding}}(tetramer) = K_{\text{interfaces}}\left(K_{folding}(monomer)\right)^4 \quad (1.1)$$

This effect accounts for the markedly enhanced resistance of the human Hb tetramer to denaturation when Hb subunits are assembly into tetramers. This resistance is strongly dependent on protein concentration, with disassembly being favored at low concentrations (Antonini and Brunori, 1971; Bunn, 1987; Hargrove et al., 1997). In erythrocytes, tetrameric Hb is highly concentrated and present at ~20 mM on a per heme basis. At these levels, tetrameric human Hb is highly resistant to dissociation into dimers and monomers, the fractions of which are extremely small. Organisms that use extracellular globins for O_2 transport have even larger molecular weights with higher degrees of oligomerization. The giant hemoglobin from *Lumbricus terrestris* (Fig. 1.5) is strongly resistant to denaturation and is assembled into a giant Hb complex composed of 144 subunits coupled to a central linker protein scaffold with a total molecular weight of ~3.6 MDa (Strand et al., 2004).

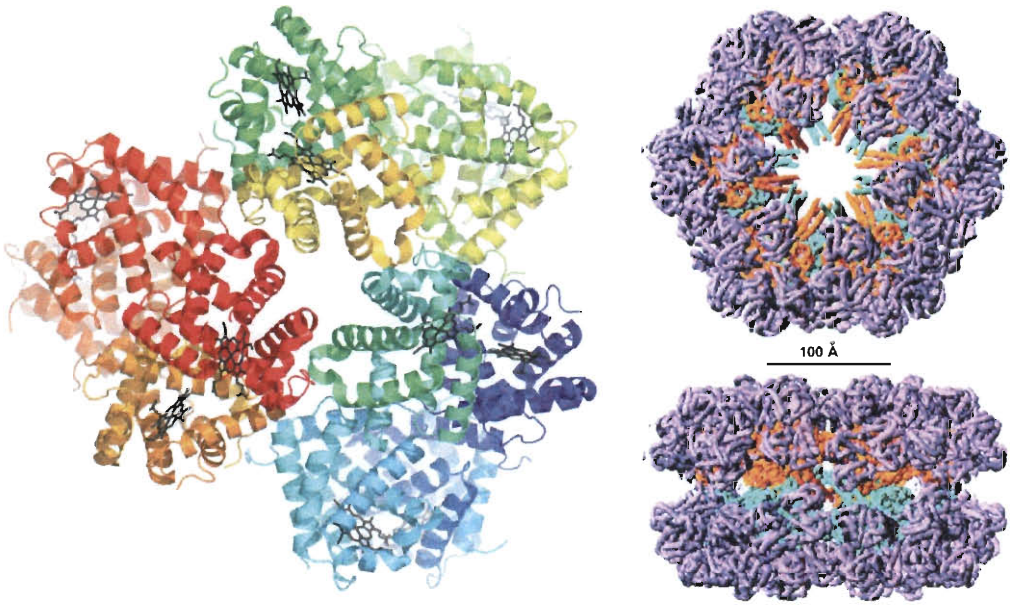


Figure 1.5. Structure of *Lumbricus terrestris* erythrocruorin. **LEFT:** one of the twelve globin dodecamers. The globins subunits are shown in cartoon and the heme cofactors are shown in dark gray sticks. Image created using PyMol and PBD entry: 1x9f. **RIGHT:** entire erythrocruorin protein with 144 globin subunits (purple) arranged around a central scaffold (cyan and orange). Image adapted from (Royer et al., 2000).

1.3. Sperm whale myoglobin: the classic globin system

The biological role of Mb

A requirement for O_2 storage and release by Mb is unclear as knock out studies in mice demonstrate normal whole animal phenotypes under basal physiological conditions. The knock-out mice show adaptation to hypoxic stress as seen by increased cardiac capillary density, coronary flow, and blood Hb content (Garry et al., 1998; Schlieper et al., 2004). In addition to the role linked to oxygen storage and release during muscle contraction, Mb has been suggested to play a role in nitric oxide (NO) scavenging (Brown, 2003; Eich et al., 1996; Flogel et al., 2001; Wittenberg and Wittenberg, 2003)

and removal of reactive oxygen species (Ordway and Garry, 2004), and these functions appear to contribute to endurance and cardiac function in mice (Merx et al., 2005). Furthermore, there appears to be a cycle in which a low oxygen tension in tissues provokes a change of Mb roles in relation to NO: in normoxic conditions Mb scavenges NO, but may help generate NO under hypoxic conditions (Rassaf et al., 2007). However, this latter role remains controversial.

For a long time it has been known that Mb content in human skeletal muscles varies according to oxygen levels (Reynafarje, 1962a, b). In addition, during periods of hypoxia-induced erythropoiesis, Mb expression has been found to be down regulated in human skeletal muscles to allow heme to be diverted to increased erythrocyte formation (Robach et al., 2007).

Pathophysiologies linked to Mb

To date, no genetic myoglobinopathies have been reported but four asymptomatic, single mutation Mb variants have been identified in humans (*i.e.*, D54K, K133N, R139Q, and R139W) (Romero-Herrera and Lehmann, 1974). The only known Mb-linked pathologies are rhabdomyolysis and myoglobinuria, which usually occur concurrently in patients that have received physical (*e.g.*, injury) or chemical (*e.g.*, alcohol or drug abuse) trauma or suffer from organ failure (Chatzizisis et al., 2008; Daher Ede et al., 2005; Knochel, 1993; Slater and Mullins, 1998). Rhabdomyolysis (rapid breakdown of muscle fibers) and myoglobinuria (presence of Mb in urine) are responsible for ~5% of acute renal failure deaths in the US (Zager, 1996). Following trauma, damaged muscle fibers can release large quantities of Mb (among other cytosolic proteins) into the bloodstream, which are normally filtered through the kidneys and eliminated. Severe trauma may lead

to renal necrosis and renal failure by obstruction of renal tubules caused by accumulation of metMb and its nephrotoxic breakdown products, which cause significant oxidative stress (Braun et al., 1970; Reeder, 2010).

Current understanding of the in vivo assembly of Mb

Synthesis of Mb involves expression of the apoMb polypeptide from a single gene and production of the heme cofactor. *In vitro*, apoMb folds properly without heme and remains relatively stable, although amyloid fibrils have been discovered after long incubation at room temperature (Fandrich et al., 2001; Hughson et al., 1990; Iannuzzi et al., 2007). In contrast, fibril formation does not occur with holoMb (Regis et al., 2005), and fibrillization of amyloidogenic apoMb (Fandrich et al., 2001) as well as induced toxicity appear to be inhibited by heme binding (Iannuzzi et al., 2007). Early interactions of heme during the initial folding of the Mb polypeptide might be a requirement *in vivo* for avoiding misfolding, and formation of amyloid fibrils at 37 °C. Despite the lack of known Mb misfolding diseases (Eakin et al., 2002), it is possible that some diseases do involve Mb aggregates, as is observed for inclusion body myositis, which occurs due to trauma or exposure to toxic chemical oxidants (Needham and Mastaglia, 2007).

Structure of apoMb

Although ultra-high resolution structures of holoMb from several species have been determined by X-ray crystallography (Birnbaum et al., 1994; Evans and Brayer, 1990; Kendrew et al., 1958; Schreiter et al., 2007; Scouloudi and Baker, 1978; Smerdon et al., 1990), the crystal structure of "native" apoMb has not been determined to date. ApoMb can be produced *in vitro* by extraction of the heme into organic solvents at low

pH (Ascoli et al., 1981). When samples of apoMb are neutralized, the resultant protein has lost about a third of its helicity as measured by its CD spectrum, but the remaining tertiary and secondary structures appear to be similar to those found in the holo-form. High resolution NMR experiments indicate loss of structure in the F helix, the EF and FG loops, the N-terminal portions of the G helix, and the C-terminus of the H helix, whereas the A, B, C, D, E, and most of the G and H helices appear intact (Eliezer and Wright, 1996; Eliezer et al., 1998; Lecomte et al., 1999). ApoMb also loses compactness as measured by small angle X-ray scattering (Nishii et al., 1994).

Equilibrium analysis of the apoMb unfolding mechanism

The unfolding of apoMb has been extensively studied *in vitro*. The existence of a folding intermediate was first proposed in 1976 (Balestrieri et al., 1976), building on previous fluorescence experiments (Kirby and Steiner, 1970) which led to the analysis of apoMb unfolding using a three-state unfolding mechanism (Barrick and Baldwin, 1993b; Hargrove et al., 1994a; Ramsay et al., 1995). The acid and urea-induced intermediates have been structurally characterized by circular dichroism (CD), site-directed mutagenesis, and hydrogen exchange NMR experiments. The results suggest that the A, G, H, and part of the B helices remain folded and docked to each other, forming a molten-globule hydrophobic core, whereas the other regions of the protein, including most of the heme pocket, are highly disordered and flexible (Eliezer et al., 2000; Eliezer et al., 1998; Griko et al., 1988; Hughson and Baldwin, 1989; Hughson et al., 1991; Hughson et al., 1990; Jennings and Wright, 1993; Nishimura et al., 2003).

The intermediate (I) has been shown to possess higher fluorescence than either the original native or unfolded states. The emission peak for this I state is at a wavelength in

between those for the folded and unfolded states when the intermediate is induced by acid or GuHCl (Ramsay et al., 1995). One theory is that partial unfolding to the I state causes an increase in fluorescence of Trp7 due to an apolar collapse of EF corner onto the G and H helices, which moves Lys79 or Met131 away from the indole ring and buries it in a more hydrophobic environment (Ballew et al., 1996; Hargrove et al., 1994a; Irace et al., 1981; Kirby and Steiner, 1970; Tcherkasskaya et al., 2000a; Tcherkasskaya et al., 2000b). Another theory is that hyperfluorescence may be caused by enhanced flexibility of the fluorophores in the intermediate state (Ervin et al., 2002).

The three-state N-to-I-to-U equilibrium-unfolding path is easily observed under acid and GuHCl denaturation conditions (Barrick and Baldwin, 1993a; Barrick et al., 1994; Griko et al., 1988; Nishii et al., 1995) but is less readily characterized when unfolding is induced by urea due to higher fluorescence of the unfolded state (Barrick and Baldwin, 1993b; Baryshnikova et al., 2005). At low pH, sw apoMb may populate two distinct intermediates (*i.e.*, I_a and I_b) that differ by additional secondary structure in helix B (Jamin and Baldwin, 1998; Loh et al., 1995); however only one equilibrium state intermediate has been reported using GuHCl and urea as denaturants at neutral pH.

The kinetics of apoMb folding

The folding and unfolding rates of apoMb have also been extensively studied. A kinetic folding intermediate of apoMb was first described in 1993 by Wright and colleagues who reported a burst phase intermediate with helices A, G, and H formed within the dead time (few ms) of the stopped-flow mixing device, which was followed by complete folding to the native apo-state several seconds later (pH 6.1, 5 °C) (Jennings and Wright, 1993). The initial submillisecond collapse to the intermediate corresponded

to an initial gain of a third of the final secondary structure (far-UV CD) and half of the change in radius of gyration (X-ray scattering) for a pH jump from 1.2 to 6 (Uzawa et al., 2004). These early microsecond CD and fluorescence changes appear to correspond to formation of the A, G, and H helix core (Ballew et al., 1996). Thus, the intermediates observed by equilibrium-unfolding and kinetic-refolding experiments are believed to possess very similar structures. The major exceptions are that the E helix, which plays no apparent role in the equilibrium intermediate, appears to stabilize the transient folding intermediate states by docking with the hydrophobic core of A, G, and H helices and that non-native interactions between theses helices are formed in the kinetic I states, which must be resolved before final folding can occur (Nishimura et al., 2006). These conclusions were also supported by the observations that the H64F and H64L mutations stabilize the I state (Smith, 2003). In most kinetic experiments, two kinetically trapped intermediates, denoted as I_a and I_b, are commonly observed and may have structures similar to the two intermediates found in steady-state experiments with urea and pH (Jamin and Baldwin, 1998; Loh et al., 1995). Thus, the refolding of apoMb may involve a three-step process with U to I_a, I_a to I_b and I_b to N phases, having half-times of approximately 5 μs, 30 ms and 0.5 s, respectively (Jamin, 2005; Jamin and Baldwin, 1998; Uzawa et al., 2004); however these kinetic intermediates are sensitive to experimental conditions, and it is difficult to establish their relevance to the equilibrium folding mechanism.

Folding of holoMb

The mechanism of holoMb unfolding has only been addressed in a few studies, and until our work, most analyses were semi-empirical and assumed an apparent two-

state equilibrium- and/or kinetic mechanism (Hargrove et al., 1994a; Hargrove and Olson, 1996; Hargrove et al., 1996b; Moczygemba et al., 2000; Ramos et al., 1999; Regis et al., 2005; Wittung-Stafshede et al., 1998). The resistance of holoMb to unfolding appears to be determined primarily by heme affinity (Hargrove et al., 1996b), and therefore, redox state and bound ligands are major factors influencing stability (Hargrove and Olson, 1996). For example, the reduced, unliganded form of Mb (deoxyMb-Fe^{II}) is much more resistant to chemical denaturation than the oxidized form (metMb-Fe^{III}), most likely due to the enhanced heme affinity (Hargrove and Olson, 1996). A quantitative analysis of the observed unfolding curves requires a detailed mechanism that accounts for the fate of the dissociated heme, *i.e.,* whether it self-associates, binds to unfolded and intermediate states, or participates in all of these processes.

The affinity of folded apoMb for heme

The heme affinity of all globins is too high to measure directly in equilibrium titrations. Estimates of the equilibrium association constant can be calculated from the ratio of heme association and dissociation rate constants, and the value for hemin binding to sw apoMb is $K_H \approx 10^{14}$ M^{-1} (Hargrove et al., 1996a). Three major factors appear to be responsible for this high affinity: (1) nonspecific hydrophobic interactions between the porphyrin and the hydrophobic pocket (10^5-10^7 M^{-1}), (2) covalent bond formation between the ferric iron and the proximal His93 (10^3-10^4 M^{-1}), and (3) specific interactions with conserved amino acids in the heme pocket (10^3-10^4 M^{-1}) (Hargrove et al., 1996b).

Rates of heme binding and release

The first kinetic studies of heme binding to apoMb (and apoHb) were carried out in the 1960s (Antonini and Gibson, 1960; Gibson and Antonini, 1960). In most cases, the reaction of free heme with apo-globin shows a biphasic time course, with a rapid bimolecular phase ($k'_H \approx 1 \times 10^8\ M^{-1}s^{-1}$) and slow first order phase, which itself shows heterogeneity but little or no dependence on apoprotein concentration. The slow phases have been attributed to the dissociation of hemin dimers and perhaps even higher-order aggregates (Adams, 1977; Asher et al., 2009; Brown et al., 1970; de Villiers et al., 2007; Hargrove et al., 1996a; Kuzelova et al., 1997), which need to dissociate before binding can occur. Using electrospray ionization mass spectrometry, Lee and coworkers (Lee et al., 1999) have suggested that hemin aggregates bind to apoMb, producing intermediates of holoMb with as many as three hemes attached at one time, although this is probably unlikely in the absence of excess hemin, and may be an artifact of the electrospray analysis.

To avoid these problems, most workers have used CO-heme, which is monodisperse up to concentrations of 10 μM due to presence of a bound ligand (Gibson and Antonini, 1960; Hargrove et al., 1996a; Light et al., 1987; Rose and Olson, 1983). Dicyano-hemin ($(CN)_2$-heme) has also been used in kinetic binding studies because it is completely monodisperse (Brown et al., 1970; Shack and Clark, 1947). However, the presence of the two axial ligands creates additional kinetic phases due to dissociation of at least one of the cyanides after binding the protein (Chiba et al., 1994; Crespin et al., 2005; Kawamura-Konishi et al., 1988).

1.4. *Cerebratulus lacteus* hemoglobin: a simpler unfolding system

Biological role of CerHb

In 1998, a mini-Hb (109 residue monomer, 12.0 kDa, Fig. 1.6.) was discovered in the neuronal tissues of the nemertean worm *Cerebratulus lacteus (Vandergon et al., 1998)*. The monomeric globin, referred to as CerHb, functions as an O_2 storage molecule, similar to Mb, for use during periods of hypoxia to anoxia when the worm is burrowing deep in sand. *In vitro,* CerHb can also bind CO and NO with high affinity (Pesce et al., 2004; Salter et al., 2008).

Figure 1.6. Overlay of sw Mb and *Cerebratulus lacteus* hemoglobin represented in gray and red ribbons, respectively. The A helix, missing in CerHb, is shown in dark gray. Figure created with PyMol using PDB entries 1kr7 (CerHb) and 1jp6 (sw Mb).

Structure of CerHb

The crystal structure of CerHb was resolved by Pesce and colleagues in 2001 (Pesce et al., 2001), and shows CerHb to be an ancestor to Mb/Hb chains (high sequence conservation); however, CerHb does not possess an A helix, and parts of the B and H helices are truncated (Fig. 1.6). The Olson/Moen group has characterized ligand binding

to CerHb in great detail, generating a library of over 150 different mutants (Pesce et al., 2004; Salter et al., 2008). The very large O_2 association rate constant is due to a hydrophobic tunnel created by the lack of the N-terminal A helix and a shortened H helix (Salter et al., 2008). Until our work, CerHb had not been characterized in terms of its folding mechanism. CerHb is of interest as it lacks the main components of the molten globule intermediate in apoMb and, therefore, might be useful as a simple monomeric globin model, in which the apoform unfolds by a two-state mechanism.

1.5. Human hemoglobin: a complex hetero-tetramer

Biological role of Hb

Vertebrate Hbs function primarily as dioxygen transporters packaged in red blood cells and deliver O_2 from the lung alveolar gas spaces to actively respiring striated muscles and neuronal tissues. The HbO_2 form of these proteins, like MbO_2, also scavenge NO by dioxygenating it to nitrate (Brown, 2003; Doherty et al., 1998; Eich et al., 1996; Flogel et al., 2001; Merx et al., 2005; Wittenberg and Wittenberg, 2003). NO produced by inducible nitric oxide synthase (iNOS) during inflammation shuts down the TCA cycle and electron transport chain of invading pathogens, but is cytotoxic to muscle cells (Gardner et al., 2000; Ullmann et al., 2004). Thus, the oxyHb present in circulating red cells actively detoxifies NO to nitrate during periods of systematic inflammation or inhalation of toxic gases.

Even though Hb has been extensively studied for over a century, little is known about the detailed mechanisms involved in Hb folding and assembly during erythropoiesis. Hb is the main component in mature erythrocytes and reaches a

concentration of ~5-5.5 mM on a per tetramer basis or 20-22 mM on a per heme basis (Messerschmidt, 2001). Among Hb variants, hemoglobin A (HbA) is the most prevalent in human adult erythrocytes (97%) and is composed of 2 α and 2 β subunits (Schechter, 2008). The β subunits found in HbA are replaced by δ and γ subunits in Hemoglobin A_2 (HbA_2) and fetal hemoglobin (HbF), respectively.

Structure of HbA

The human adult HbA is composed of 2 α and 2 β subunits of 141 and 146 residues, respectively, that assemble into a hetero-tetramer with two strong interfaces ($\alpha_1\beta_1$ and $\alpha_2\beta_2$) and two weak interfaces ($\alpha_1\beta_2$ and $\alpha_2\beta_1$) (Fig. 1.7). The subunits when assembled into the tetramer possess the characteristic 3-on-3 helix fold (Fig. 1.2), with 7 and 8 helices present in the α and β subunits respectively, the α subunit lacking the D helix. Human HbA exists in two major conformational states: relaxed (R) and tense (T) states. The R state has a higher O_2 affinity, and in the absence of O_2, the R state reverts to the more stable T state. The inter-subunit interactions influence O_2 binding and vice versa resulting in cooperativity.

Erythropoiesis

Erythropoiesis is the development of mature red blood cells. Erythrocytes (red blood cells) are the primary oxygen cellular carriers of O_2 in our body and have been the focus of physiological studies since William Harvey's description of the human circulating systems in the 1600s (Ebert and Bunn, 1999; Ferreira et al., 2005; Orkin, 1995; Pang et al., 2005; Zhu and Emerson, 2002). Mature erythrocytes of healthy adults are formed from hematopoietic stem cells in the red marrow of large bones in about

seven days and are released into circulating blood at a rate of ~2.4 million per second (Sackmann, 1995). The maturation process is very complex and involves the progressive loss of organelles (nucleus, mitochondria, lysosomes, and ribosomes), accompanied by production of specific proteins, mostly Hb as well as erythropoietin, transferrin receptors, glycophorin and spectrin (Ebert and Bunn, 1999; Ferreira et al., 2005; Orkin, 1995; Pang et al., 2005; Zhu and Emerson, 2002).

HbA is the major component of mature adult erythrocytes. The synthesis of Hb is governed by many mechanisms that ensure that heme, α-, and β-chains are synthesized in the correct 4:2:2 ratio, because any imbalance can have severe toxic effects on the red cells (Chen, 2007). The levels of heme are critical during erythropoiesis as they also regulate gene transcription and translation of the Hb chains (Chen, 2007). For example, heme has been shown to regulate the transcriptional factor Bach1 during erythroid differentiation (Igarashi and Sun, 2006; Taketani, 2005) as well as translation in erythroid precursors by modulating eIF2α kinase activity of heme-regulated inhibitor kinase (HRI) to reduce excessive production of globins (Chen, 2007). It has been suggested that mammalian cells might contain a pool of free heme, which could serve both precursor and regulatory functions (Ponka, 1999; Wagener et al., 2001). Interestingly, because the heme binding affinities of Bach1, heme regulated inhibitor (HRI) and heme oxygenase (HO-1) appear to be on the order of 10^{6-7} M^{-1} (Chen, 2007), the levels of free heme in erythrocytes are most likely in the high nanomolar to low micromolar range.

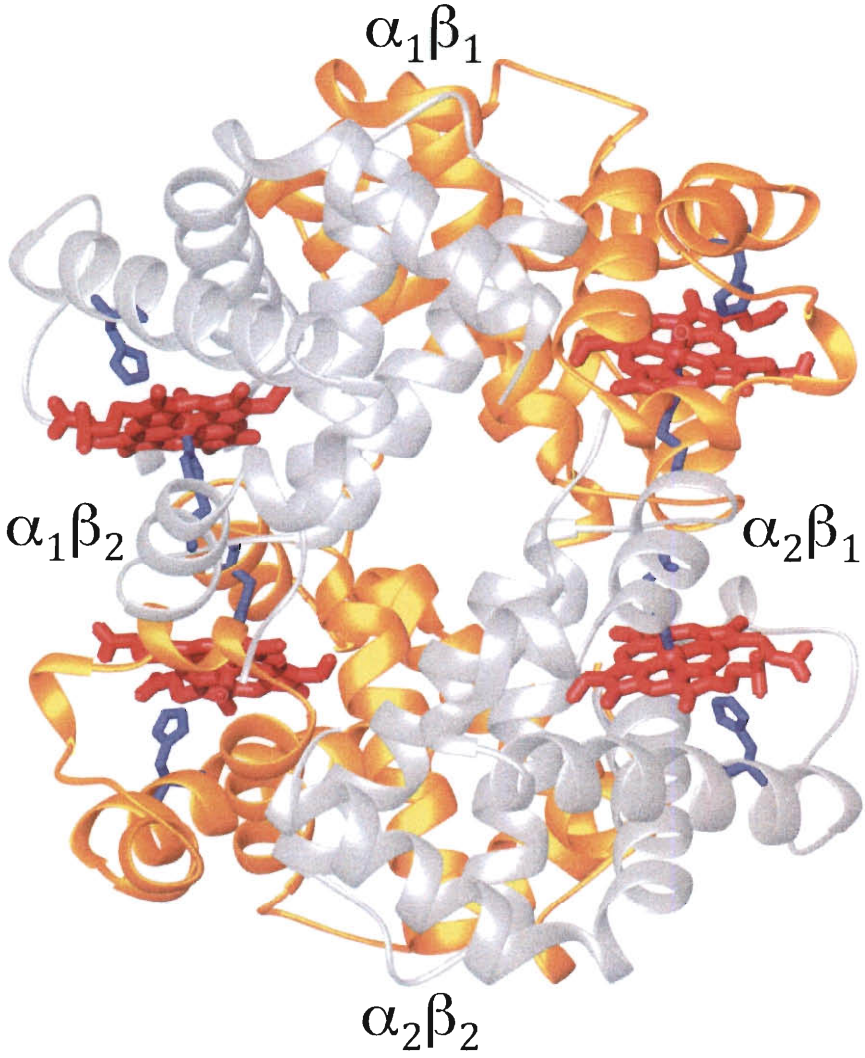


Figure 1.7. Crystal structure of adult human oxyHbA. The 2 α and 2 β chains are represented in silver and gold ribbons, respectively. The four hemes are shown in red sticks bound to the respective proximal and distal histidines in blue. Image created using PyMol and PDB entry: 2dn1.

Erythropoiesis, which involves high level of Hb expression in erythroid precursors, is also dependent on the availability of large amounts of heme to avoid anemias (*i.e.*, at least 20 mM levels bound to Hb). Interestingly, heme in erythrocytes is produced at a rate dependent on the insertion of iron into porphyrin rings by ferrochelatase, contrary to non-erythroid cells where the formation of δ-aminolevulinic

acid is the rate limiting step in heme biosynthesis (Ponka, 1997). The iron requirement in erythroid cells is solely provided by transferrin (Hemmaplardh and Morgan, 1974a, b). In addition to heme-linked regulation of erythropoiesis, erythropoietin serves as a feedback loop to control the amount of synthesized red blood cells according to oxygen demand and red blood cell destruction rate (Socolovsky, 2007). In addition, the recently identified peptide-hormone hepcidin has been suggested to play a role in the regulation of iron recycling and iron balance (Ganz, 2006), which in turn affects Hb production and erythropoiesis (Nicolas et al., 2002).

Hemoglobinopathies

Hemoglobin disorders, or hemoglobinopathies, occur due to defects in globin structures or gene expression. Any mutation that leads to a change in secondary or tertiary structure in one of its subunits can be the cause of pathophysiologies. Among several diseases linked to structural changes in Hb, sickle cell anemia is the best characterized medically and is due to a single-point mutation (E6V) in the β-subunit, which induces a conformation change of the tetramer ($\alpha_2\beta^S_2$) and abnormal erythrocyte morphology (Bunn and Forget, 1986; Schechter, 2008).

Synthesis of Hb is strictly regulated and any deviation from the 2:2:4 ratio of α:β:heme can lead to severe cytotoxic effects (Chen, 2007). Hb in healthy human adults is located exclusively in mature erythrocytes circulating in blood plasma where it accounts for two thirds of total iron in humans (Stojiljkovic and Perkins-Balding, 2002). However, a small amount (5 to 50 mg/l) of Hb is normally present in plasma as a result of spontaneous hemolysis and is normally transported back to the liver as dimers bound to haptoglobin (Stojiljkovic and Perkins-Balding, 2002).

Thalassemias are characterized by an inherited imbalance in expression of HbA subunits and can cause anemia-like symptoms in humans (Bank, 2007; Schechter, 2008; Yu et al., 2007). The α- and β-thalassemias have distinct pathologies that are directly linked to which subunit fails to be produced in the proper amount. In α-thalassemia, the β-chains are being produced in excess, which usually produces a less severe anemia than in β-thalassemia where the α-chain is expressed in excess. This difference is due to two factors: first, α-subunits are encoded by two genes (four alleles) whereas only one gene (two alleles) encodes the β-chains, so that loss of one α-gene is not as severe as the loss of one β-gene; second, the β-subunits of Hb are more stable than α-subunits and can reversibly form β_4 tetramers (HbH) (Steensma et al., 2005). The excess α-chains in β-thalassemias leads to toxic aggregates that accumulate in cell membranes as well as in red cell inclusion precipitates called Heinz bodies (Bank, 2007).

Importance of heme homeostasis

Heme is present in almost all living organisms, except for a few anaerobes and a few unicellular organisms auxotrophic for heme (and related porphyrins) (Ponka, 1997). Heme biosynthesis in mammalian cells is a finely tuned process, which consists of a well-conserved eight-enzyme pathway, located in the cytoplasm and mitochondria, with final formation of heme occurring in the inner-membrane of mitochondria. The heme must then be transported through the mitochondrial membranes to help form functional cytosolic heme-proteins including Mb and Hb (Hamza, 2006). The levels of heme are tightly regulated at several control points in the anabolic and catabolic pathways (Bunn and Forget, 1986; Hamza, 2006; Jeney et al., 2002; Ponka, 1999; Raju and Maines, 1996;

Wagener et al., 2001). In turn, these levels regulate gene transcription and translation of Hb chains (Chen, 2007; Han et al., 2001; Taketani, 2005).

Human plasma contains many heme-binding proteins (*i.e.*, albumin, hemopexin, and lipoproteins) that deliver heme to the liver (Piccard et al., 2007; Stojiljkovic and Perkins-Balding, 2002). Heme released from heme-proteins can readily generate oxygen radicals and catalyze the oxidation of lipids, proteins, and DNA (Graca-Souza et al., 2002). Heme released from Hb after episodes of intravascular hemolysis has been suggested to contribute to acute renal failure (Graca-Souza et al., 2002). Heme has also been considered a causative agent in organ failure after ischemia-reperfusion, in part because heme-oxygenase (enzyme which breaks down heme) induction in heart and kidney tissue is slow, allowing an accumulation of heme in cell membranes and lipoproteins which has been shown to produce large amounts of radicals and reactive oxygen species (Graca-Souza et al., 2002; Jeney et al., 2002; Raju and Maines, 1996; Wagener et al., 2001).

Heme is a very good source of iron for humans and bacteria. Some micro-organisms are unable to synthesize heme and require large amounts of heme to grow *in vitro*. The availability of heme plays a role in virulence of certain bacteria (Stojiljkovic and Perkins-Balding, 2002) in part because most pathogens prefer heme as an iron source (Rouault, 2004). Consequently, the accumulation of free heme in cells and in blood is tightly regulated and suppressed to avoid pathogenesis (Rouault, 2004).

Human apoHbA dimers

Removal of heme by extraction into acid-acetone mixtures and then refolding by slow dialysis into neutral pH buffer (Ascoli et al., 1981) leads to formation of $\alpha_1\beta_1$ apo-

dimers (Fanelli et al., 1958; Rossi-Fanelli et al., 1964; Winterhalter and Huehns, 1964; Yip et al., 1972), which have 48% helicity of native holoHbA (Hrkal and Vodrazka, 1967). These apoHb dimers are still capable of reacting rapidly with free heme to reconstitute intact and functional holo-tetramers (Winterhalter and Colosimo, 1971; Winterhalter and Huehns, 1964). When heme is removed from isolated α and β subunits, almost all secondary structure is lost; the apoproteins precipitate at temperatures greater than 15 °C; and reconstitution with added heme does not occur in the absence of partner subunits (Waks et al., 1973).

Heme binding to apoHb

The binding of heme to folded apoHb has been studied for ~50 years (Adams, 1976; Antonini and Brunori, 1971; Antonini and Gibson, 1960; Gibson and Antonini, 1960) and most studies have shown that heme binds very tightly (K_d~ 10^{-11} to 10^{-14} M) to both apoMb and the α and β subunits of apoHb. The α subunit has a higher affinity for hemin than the β subunit (Cassoly and Banerjee, 1971; Winterhalter et al., 1968) despite similar association rate constants for heme binding to the two different apoglobin subunits (Rose and Olson, 1983). Thus, there is a redistribution of heme from the β to the α subunit over time in experiments involving one heme per apoHb dimer (Winterhalter et al., 1968; Winterhalter and Deranleau, 1967). This difference in affinity was confirmed in hemin dissociation experiments which showed that the rate of hemin dissociation from the β subunits is ~3 fold larger than from the α subunits in metHb tetramers and dimers (Hargrove et al., 1997). In addition, hemin affinity differs between the monomers, dimers and tetramer (Hargrove et al., 1997) and losses of heme from the Hb subunits causes

much larger loss of secondary structure than what is observed during apoMb formation (Leutzinger and Beychok, 1981).

How is Hb synthesized?

The α and β subunits chains of Hb are formed separately on the ribosome whereas heme synthesis mostly occurs within mitochondria. It is still uncertain how the Hb chains fold (co-translationally or post-translationally) and when heme specifically interacts with the nascent chains. Most workers believe that apo-globins first fold and then bind heme; however, heme could interact prior to a fully formed apo-globin (Komar et al., 1993). Mammalian Hb chains are translated in about one minute (Dintzis, 1961; Goustin and Wilt, 1982), whereas folding of globin chains *in vitro* is accomplished in a few seconds. Heme binding to the folded state also occurs very rapidly (Hargrove et al., 1996a; Rose and Olson, 1983). *In vitro* studies have demonstrated that without heme present, globin synthesis does not proceed, and those apo-chains that are made are unstable and rapidly aggregate (Chen, 2007).

Until recently, hemoglobins and myoglobins were believed to fold without the help of molecular chaperones. The discovery of the α-hemoglobin stabilizing protein, AHSP, by Weiss and coworkers (Gell et al., 2002; Kihm et al., 2002) has refocused interest on the assembly of Hb. Even though free α-chains are cytotoxic by producing precipitates and reactive oxygen species (Feng et al., 2004; Feng et al., 2005; Scott et al., 1993), a small excess of α-chains are normally formed in pre-erythroid cells (two alleles for α versus one allele for β chains) forming free pools of α-chains (Baglioni, 1968; Baglioni and Campana, 1967; Shaeffer, 1967; Tavill et al., 1967). AHSP is thought to be

an erythroid chaperone that stabilizes free α-chains in the absence of β-chains (Bank, 2007; Kihm et al., 2002; Weiss et al., 2005; Yu et al., 2007). In addition, heme delivery to nascent chains has been suggested to be the limiting factor in formation of functional Hb (Levere and Granick, 1965). Thus, AHSP may participate in helping proper assembly of Hb by assuring that α chains are not degraded prior to heme incorporation (Mollan et al., 2010).

Hemoglobin assembly during erythropoiesis is extremely complex and the question of when heme binds to globins chains still remains unresolved despite many efforts in the hematological, biochemical, and biophysical fields. Because apoHb dimers form when heme is removed from tetramers, even at high concentrations, we assume that the α and β subunits first fold up into apo-dimers which then bind hemin. These ideas are explored in Chapter 7.

1.6. General relevance and heme protein biology

Biophysical

More than 30% of proteins in cells bind a cofactor in order to function properly (Gray, 2003). Most, if not all, cofactor-binding proteins are significantly more stable in their holo form compared to their apo-form, as reviewed in (Wilson et al., 2004). This situation is clearly the case for Mb (Hargrove and Olson, 1996), azurin (Leckner et al., 1997), and flavodoxin (Apiyo and Wittung-Stafshede, 2002) (Fig. 1.8).

Figure 1.8. Crystal structures of three model cofactor-proteins. Ribbon representation of *D. vulgaris* flavodoxin (PDB entry: 2fx2) with flavin mononucleotide (FMN, purple sticks), *P. aeroginosa* azurin (PDB entry: 1azu) with copper ion (blue sphere), and sw Mb (PDB entry: 1jp6) with heme (red sticks with the iron as a sphere). Images created using Pymol.

The folding and assembly of cofactor-binding proteins is in most cases poorly understood. Some proteins have been suggested to retain their cofactor upon unfolding, although the binding is normally much weaker (*e.g.*, cytochrome b_{562} (Robinson et al., 1997) and azurin (Marks et al., 2004)). Heme has been suggested to bind weakly to the unfolded Mb polypeptide after denaturation (Hargrove and Olson, 1996). Heme might also be able to bind more specifically to folding intermediates and modulate the kinetic and equilibrium unfolding/refolding pathways (Fig. 1.9).

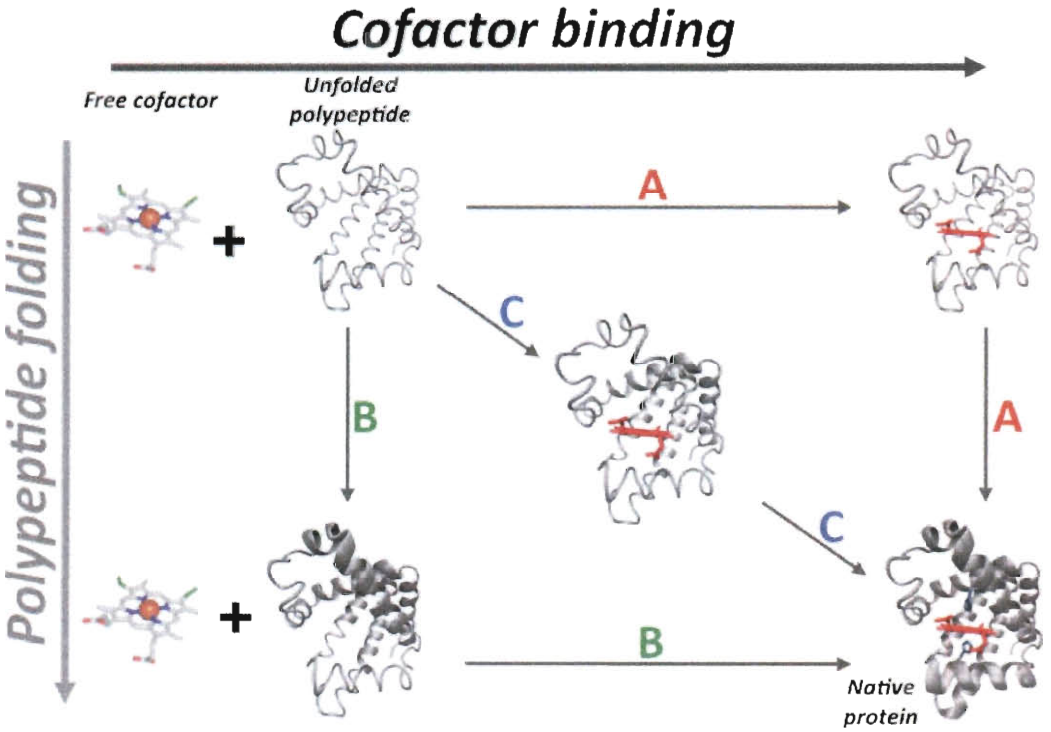


Figure 1.9. Three possible assembly mechanisms involved in Mb assembly, which includes polypeptide folding and cofactor binding. The mechanisms are represented where (**A**) the cofactor binds first, (**B**) the cofactor binds after folding, and (**C**) the cofactor binds simultaneously with folding. Ribbon representation of the crystal structure of WT aquomet sw Mb (PDB entry: 1jp6) using Pymol. The NMR structures for the apoglobin native (Eliezer and Wright, 1996; Eliezer et al., 1998; Lecomte et al., 1999) and intermediates (Eliezer et al., 2000; Eliezer et al., 1998; Griko et al., 1988; Hughson and Baldwin, 1989; Hughson et al., 1991; Hughson et al., 1990; Jennings and Wright, 1993; Nishimura et al., 2003) states are represented in ribbon and thin loops for the folded and unfolded regions, respectively. Figure adapted from Wilson et al. (2004)

Understanding of the role of heme in the unfolding of Mb and Hb *in vitro* will establish new insights in the folding pathways of homologous heme-proteins (*e.g.*, flavoHbs, neuroglobins, and cytoglobin). In addition, our new work was designed to test quantatively whether heme can bind only after the apoglobin chains are completely folded or if it can bind intermediate states (Bank, 2007; Hargrove et al., 1996a; Yu et al., 2007).

Clinical/pathophysiological conditions

All the factors involved in the formation of Hb are relevant in understanding erythropoiesis. Unraveling how heme modulates the folding of Mb, and Hb, will lead to a better understanding of the synthesis and degradation of Hb, and pathologies associated with unstable mutants. These results may provide insights to control anemias caused by unstable hemoglobins and the equally serious thalassemias, caused by unbalanced expression of the α and β chains.

Finally, understanding the factors that contribute to Mb denaturation may lead to the design of new treatments to reduce toxic accumulation of apoMb and hemichrome intermediates in renal tubules of patients suffering from rhabdomyolysis (in combination with various continuous hemofiltration methods) (Bellomo et al., 1991; Naka et al., 2005). Myocyte apoptosis induced by oxidation and hemin loss from Mb has been shown to play a role in ischemia-reperfusion injury (Scarabelli and Gottlieb, 2004), and thus understanding of how heme modulates the *in vitro* formation of Mb may allow us to inhibit the rate and extent of Mb unfolding, which could help treat injuries due to muscle trauma and ischemia/reperfusion.

Biotechnological application

One of the medical dilemmas of today is the widespread shortage of blood available for blood transfusions required for immediate treatments of unremitting military incidents and civilian disasters (Olson, 2005). Therefore, the creation of a non-pathogenic blood substitute that is commercially feasible is extremely important. One approach is to engineer a blood substitute using recombinant Hb over-expressed in *E. coli* (Dou et al., 2002; Graves et al., 2008). However, low expression yields and prohibitive costs have

precluded development of Hb products even though the major side effects such as high blood pressure, vascular lesions, and gastrointestinal dysmotility have been inhibited by re-engineering the active site of the 2 α and β subunits (Dou et al., 2002; Olson, 2005).

Mb has been used as a model system for designing heme protein based blood substitutes for several years. One of the factors that limits the expression yield of holoMb and holoHb in *E. coli* appears to be correlated with the stability of the apoglobins (Dou et al., 2002; Graves et al., 2008; Smith, 2003). In addition, production of intact holoMb and holoHb in *E. coli* is limited by the rate of heme uptake (Olson, 2005). The understanding of how the stability of Mb in the folding pathway is modulated by the presence of heme and determining when heme binds will allow the design of recombinant proteins that are more stable and express at higher levels. This information could be vital for optimization of recombinant Hb production and reduction in costs. Exogenous heme may need to be added or heme transporters co-expressed at the time of induction to avoid possible aggregation products formed from apoHb subunits.

1.7. Thesis goals and organization

The overall goal of this thesis was to investigate the role of heme in the unfolding and assembly of globins. The work and interpretations are divided into six chapters (Chapters 3-8), followed by references, and an appendix that includes the mathematical derivations for the equations in Chapters 3-7 and computer code (Chapter 5). Materials and Methods are given in Chapter 2.

Chapter 3 presents the theoretical background for resolving the effect of heme remaining non-specifically attached to the unfolded polypeptide, and self-associating to

form free hemin dimers. Remarkably, no one had derived expressions for quantifying how heme binding modulates the folding of hemoproteins. Three possible unfolding mechanisms were postulated and used to simulate the dependence of protein concentration on the apparent stability of a monomeric globin that unfolds by a simple two-state mechanism.

In Chapter 4, a detailed theory for the unfolding of holoMb is presented and takes into account the intermediate observed in the unfolding of the apoglobin. Sw Mb was chosen as the classic monomeric globin to investigate whether heme binds only to the fully folded N apo-state or also binds to the apoMb folding intermediate. Sw Mb is homologous to human Mb with 82% sequence identity; however its increased stability in both the apo and holo forms allows for mutations to be made without complete loss of structure and expression in *E. coli*. I used a mechanistic approach in which the equilibrium-unfolding of the holoprotein can be interpreted in a six-state scheme, based on the three-unfolding states N, I, and U of apoMb. The use of mutagenesis to markedly change the stability of the N apo state or its affinity for hemin was crucial for verifying our mechanism and interpretations.

In Chapter 5, the effects of hemin dimerization and protein concentration, on the unfolding of holoMb were investigated using the six-state scheme from Chapter 4 and a hemin dimerization step, which, as shown in Chapter 3, can affect the observed unfolding curves but markedly complicates the analysis. Unfolding curves of WT and H97D holoMb were measured at three different protein concentrations and then fit to the six-state scheme with hemin dimerization.

In Chapter 6, *Cerebratulus lacteus* hemoglobin was used as a simple monomeric model to verify our approaches and analyses. CerHb does not appear to populate an intermediate in the apoform and thus we could use a four-state scheme to analyze the equilibrium unfolding of holoCerHb.

In Chapter 7, my goals were two fold: first, to build equilibrium-unfolding models to explain how the apoHb dimer assembles, and second, to prove that hemichrome formation also occurs during the unfolding of aquomet HbA. Even though the results in this chapter are limited, they provide evidence that a hemichrome is on pathway for the assembly of HbA, and such hemichromes are directly relevant to the biology of the eythrocyte and formation of Heinz bodies. My thesis conclusions are presented in Chapter 8 and include correlations between chemical denaturant and thermal stabilities for apo and hologlobin variants, and provide insights and directions for future experimental projects.

Chapter 2

Materials and methods

2.1. Protein preparation

Sperm whale Mb

Wild-type sperm whale myoglobin was prepared according to Springer and Sligar (1987) as modified by Carver *et al.* (Carver et al., 1992). The H64F and V68T mutant genes were constructed using the cassette mutagenesis system developed by Barry A. Springer and Karen Egeberg at the University of Illinois (Springer et al., 1989). The T67P and H97D mutants were reconstructed at Rice University for this work using oligonucleotide–directed mutagenesis with a pET29 vector containing the gene for WT Mb to increase yields. All variants were expressed and purified at Rice University following the procedures used for WT Mb, except for the T67P and H97D mutants, which were expressed in *E. coli* BL21-DE3 (Stratagene) and grown at 30 ºC for 16-20 hrs post lag phase in Luria-Bertani media with 50 μg/mL kanamycin (Sigma). If needed, the protein was purified further using a Superdex-200 gel filtration column (Amersham Biosciences) attached to an FPLC (Amersham Bio AKTA), and the purity was ≥ 95% as assessed by SDS-PAGE gels and Soret/280 absorbance ratios. The holoMb concentrations were determined spectroscopically using ε_{409}= 157 mM^{-1} cm^{-1} for the aquomet forms of WT, T67P, V68T, and H97D Mb (Antonini and Brunori, 1971) and with ε_{393}= 93 mM^{-1} cm^{-1} for the ferric pentacoordinate H64F variant (based upon the CO-

bound form at 424 nm using ε_{424}= 187 mM^{-1} cm^{-1} for all the MbCO samples (Antonini and Brunori, 1971).

Cerebratulus lacteus Hb

Recombinant wild-type CerHb was expressed and purified at Rice University, using a synthetic gene and codon usage optimized for expression in *E. coli* (Pesce et al., 2004; Pesce et al., 2002). Reduced protein was oxidized with potassium ferricyanide, and isolated with a G-25 gel-filtration column with 10 mM phoshate pH 7. The protein concentration was determined using ε_{406}= 179 mM^{-1} cm^{-1}.

Human adult HbA

Native Human Hb was obtained from outdated blood units from the Gulf Coast Regional Blood Bank in Houston and was prepared according to (Reisberg, 1980). Briefly, the protein was isolated by centrifugation and lysis of the erythrocytes, purified by gel chromatography, and stored in the CO form at -80 °C. The aquomet form used in the experiments was obtained by replacing CO with O_2 under an intense light, and addition of a slight excess of potassium ferricyanide to oxidize the heme iron. Excess oxidizing agent was immediately removed by G25 gel chromatography. The oxidized protein was used immediately to avoid heme loss and degradation. The protein concentration was determined spectroscopically using ε_{406}= 179 mM^{-1} cm^{-1} based on the CO form (Antonini and Brunori, 1971).

Preparation of apoproteins

ApoMb and apoCerHb were prepared using a methyl-ethyl ketone extraction method at low pH (Ascoli et al., 1981; Hargrove et al., 1994b). I was able to optimize the preparation of apoMb by ensuring that the protein was dilute (> 100 μM), and only contained trace amounts of salt and buffer prior to the addition of acid. We were able to obtain yields of >90% apoprotein from wild-type Mb. The resultant apoprotein was then filtered to remove any precipitated protein and used immediately. The concentrations of apoMb and apoCerHb variants were determined spectroscopically using ε_{280}= 15.2 (Light, 1987) and 12.5 mM^{-1} cm^{-1} (estimation from the aromatic residue content), respectively. ApoHbA was prepared using the acid-acetone method (Ascoli et al., 1981) and protein concentration was determined using ε_{280}= 12.7 mM^{-1} cm^{-1} on a subunit basis. Residual heme was very minimal in all apoproteins with less than 0.5% for apoMb and apoCerHb, and less than 2% for apoHbA as determined by the Soret to 280 nm absorbance ratios.

2.2. Theoretical and experimental approaches to protein folding

In most cases, proteins fold spontaneously. From an energetic perspective, this means that the folding process is thermodynamically favorable and the folded conformation is in a state of lower free energy compared to the unfolded state (or ensemble). The difference, expressed as the Gibbs free energy change ΔG, is a measure of a protein's stability, and is governed by the balance between enthalpic (ΔH) and entropic (ΔS) contributions depending on the temperature (T), as represented by:

$$\Delta G = \Delta H - T\Delta S \qquad (2.1)$$

Upon folding, proteins form hydrogen bonds, salt bridges, and hydrophobic associations overcome the entropic cost of restricting the number of random structures. The folding free energy of most proteins is only ~ 5-10 kcal/mol, which is equivalent to a few hydrogen bonds, tipping the balance in favor of the folded state. The folding process can be mathematically simulated using a funnel-like energetic landscape, where the protein is allowed to fold to the state of lowest energy through a large number of pathways and intermediates. *In vitro,* conversion of folded to unfolded states of proteins can be induced upon addition of chemical denaturants or heat, and the process can be monitored experimentally by several methods. As proteins increase in size (and length), the folding process becomes more complex, often involving intermediates and requiring chaperones, which renders quantitative analyses more difficult.

2.3. Spectroscopic characterization of protein folding

Proteins possess unique primary, secondary, tertiary, and quaternary structures, which can be characterized in solution by various spectroscopic methods. Spectroscopy relies on how magnetic fields and electro-magnetic radiation can be altered in the presence of proteins. The main spectroscopic methods used for biophysical characterizations of proteins are: Electron Paramagnetic Resonance, X-ray crystallography, Infra-Red, Raman, UV-Vis absorption, Fluorescence, Circular Dichroism (CD), and Nuclear Magnetic Resonance. In this work, UV-Vis absorption, far-UV CD, and fluorescence emission methods were used to monitor the unfolding of apo- and holo-globins.

Fluorescence

Fluorescence is the rapid emission of light by a fluorophore that has been excited by absorbed light. Emission occurs when an orbital electron relaxes from the excited state S_1 to its ground state S_0 by emitting a photon of longer wavelength than the excitation light. The bathochromic shift between excitation and emission is caused by the loss of energy through non-radiative relaxations, releasing heat to the solvent, upon excitation.

Fluorescence can be used to study protein folding as a non-damaging technique to monitor the change of environment of either intrinsic (*e.g.*, Trp, Tyr, or Phe) or extrinsic (fluorescent labels/dyes) fluorophores of a protein during folding/unfolding. For example, the emission from a Trp residue is very sensitive to its environment and can therefore be used as a tool to investigate tertiary structure empirically.

Circular dichroism

Circular dichroism relies on the differential absorption of left and right circularly polarized light upon interaction with a protein. The spectrum in the far-UV is dominated by the amide chromophore, which is diagnostic of the backbone conformation (secondary structure), whereas the spectrum in the near-UV provides information on the aromatic and disulfide transitions, which are dependent on the side-chain environment (tertiary structure). In the far-UV range (190-250 nm), the spectrum can reveal important components of a protein's secondary structure: α-helix, β-sheet, β-turn, and random coil. The characteristic 208-222 nm and 215 nm minima for proteins with high α-helical and β-sheet content, respectively, are due mainly to the $n \rightarrow \pi^*$ transition and static-field mixing between the $n \rightarrow \pi^*$ and $\pi_0 \rightarrow \pi^*$ transitions (Berova et al., 2000).

UV-visible absorbance

Absorbance in the UV-visible region is an important quantitative tool for characterization of the electronic transitions of transition ion metals and highly conjugated organic compounds. Absorbance in the visible region gives rise to the observed colors of these cofactors, making it a very useful assay for rapid identification of these compounds. In the case of globins, changes in the axial ligand and oxidation state cause large changes in the cofactor's UV-visible absorption spectrum, which can be used to monitor various processes including unfolding and heme loss.

Equilibrium-unfolding

Measurements of equilibrium-unfolding are used to investigate the stability of proteins. Unfolding is triggered by addition of denaturant and monitored by following changes in secondary and tertiary structures using CD and/or Trp or Tyr fluorescence measurements, respectively. Sets of samples are prepared using various amounts of denaturant along with appropriate amounts of water, buffer, and protein. The presence of denaturant (*e.g.*, acid, GuHCl, GuSCN, or urea) is believed to destabilize the native folded state of a protein with respect to its unfolded state by strong electrostatic interactions with the highly polar amide backbone. To ensure equilibrium, a minimum incubation time must be determined after which the interconversion rates between folded and unfolded states are equal. The ability of a protein to fold back to its original state once unfolded is termed reversibility, and is important for accurately determining the stability and folding constants of a protein using equilibrium analyses. Information concerning the composition of each sample (*i.e.*, folded, intermediate, and unfolded states) can then be retrieved to determine thermodynamic stability parameters. In simple

systems, proteins unfold by a two-state, highly cooperative process that is well defined and resembles a phase transition.

2.4. Protein experiments

Chemical equilibrium-unfolding

The samples were prepared using 8 M stock guanidine hydrochloride (GuHCl) of very high purity (Sigma), appropriate amounts of potassium phosphate buffer (Fisher Scientific) to bring the concentration to 10 or 100 mM at pH 7, purified milliQ water and concentrated protein. The mixtures were left to equilibrate for two hours at the desired temperature. A Jasco 810 was used for CD experiments, a Varian Cary Eclipse for fluorescence emission, and a Varian Cary 50 for UV/Visible absorbance measurements. Each sample chamber was equipped with a Peltier temperature regulator. A reference hemichrome spectrum was generated by adding high concentrations of imidazole to the ferric myoglobins using a pH-adjusted 1 M stock solution from the powder form (Sigma). To examine the reversibility of hemichrome formation, metMb samples containing hemichrome intermediates were rapidly reduced anaerobically by addition of sodium dithionite (Fluka).

Thermal unfolding

Thermal melts monitored by CD at 222 nm (secondary structure) were performed on a Jasco 810 spectropolarimeter equipped with a Peltier unit and unfolding was performed in a 1 mm quartz cuvette with a scan rate of 1 °C/min for both the apo and holo proteins.

Hemin dissociation assays

The measurements of the rates of hemin dissociation were carried out using 4 μM of the desired metglobin with 40 μM apo H64Y/V68F sw Mb (termed hemin loss reagent) as described in (Hargrove et al., 1994b). The heme loss reagent in the ferric form has a different absorbance spectrum in the visible range than most ferric globins, which allows for the transfer of heme between the desired metglobin and the hemin loss reagent in the apo form to be followed spectrally. The globin of interest must be in the ferric aquomet form, as heme loss in the reduced state is too slow to measure (Hargrove et al., 1996b, k_{-H}> several days). Globins at equilibrium continuously lose and rebind heme, but in the presence of the apo-reagent in excess (10-fold in our experiments), most of the lost heme is immediately scavenged by the apo-reagent, generating the "green" holoMb reagent.

Chapter 3

General mechanisms for the unfolding of monomeric globins

This chapter is adapted from:

Culbertson, D.S., and Olson, J.S. (2010). Folding and Assembly of Myoglobins and Hemoglobins. In Protein Folding and Metal Ions: Mechanisms, Biology, and Disease, P. Wittung-Stafshede, and C.M. Gomes, eds. (Taylor and Francis, Inc.) pp. 97-122.

3.1. Introduction

Monomeric holo-globins and related type *b* heme monomers (*e.g.*, cytochrome b_{562}) appear to unfold at equilibrium by simple two-state mechanisms at much lower pH values or higher denaturant concentrations than the corresponding apoproteins. In all cases, the holoprotein is stabilized markedly by the presence of heme (Hargrove and Olson, 1996; Robinson et al., 1997). There have been many speculations on the fate of heme during the unfolding of globins. The observed lack of protein concentration dependence during chemical (Hargrove and Olson, 1996) and thermal (Robinson et al., 1998) denaturation of holo-Mb and cytochrome b_{562}, respectively, suggests that heme may remain bound even after unfolding of these hemoproteins. To examine the possible causes for the lack of protein concentration dependence, I evaluated three possible equilibrium-unfolding mechanisms for a monomeric holoHb (NH, where N represents the native fold with heme, H, bound to it). The models include: *(1)* no hemin (H) binding to the unfolded apoHb state (U); *(2)* binding of hemin to the U state (UH); *(3)* and self-

association of hemin to form dimers (H_2). To simplify the analysis and focus on the fate of the heme, the apo-form of the monomeric globin was assumed not to possess an intermediate in the folding pathway of the apoglobin. Mathematical expressions were derived to simulate the dependence of the unfolding curves on protein concentration for a simple monomeric globin. These simple mechanisms differ significantly with respect to the dependence of the unfolding curves on the absolute initial holoprotein concentration, and provide insights that were further tested in Chapters 4 and 5.

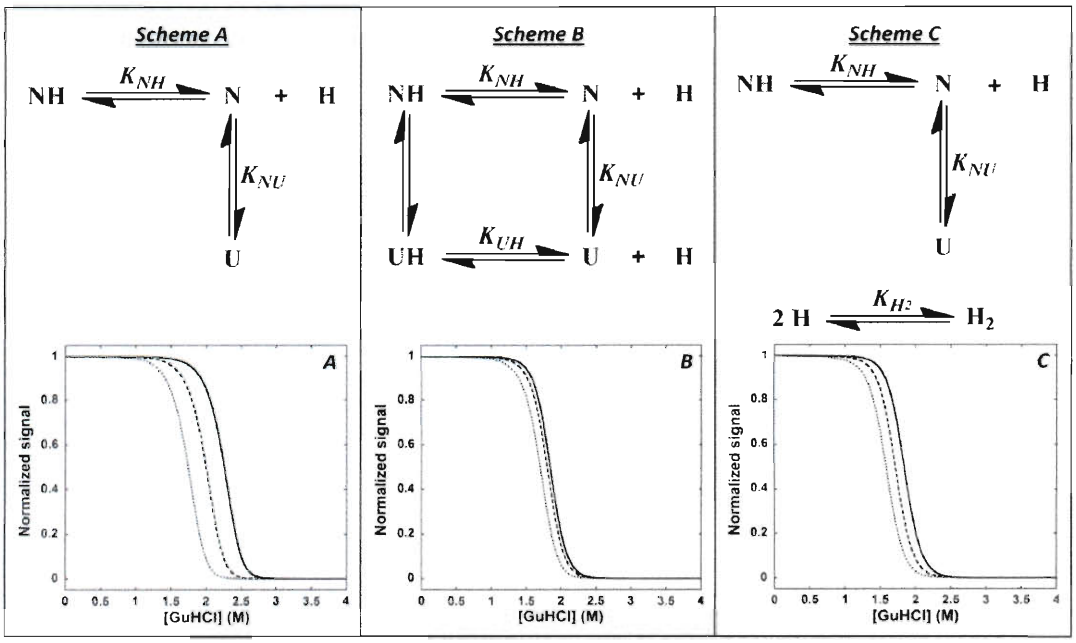


Figure 3.1. Simulated effect of protein concentration on the unfolding curves of monomeric globins for three different mechanisms with different fates of dissociated hemin. The panels on the left describe the reaction schemes, and the panels on the right show simulated unfolding curves at P_0 = 1 μM (dotted line), 20 μM (dashed line) and 400 μM (solid line). **A.** Simple mechanism where hemin remains free and monomeric after dissociation and protein unfolding (Appendix A.1.). The following constants were used: K_{NU}^0 = 1.0 x10^{-3}, m_{NU} = 3.1 kcal mol^{-1} M^{-1}, K_{NH}^0 = 10^{-12} M, m_{NH} = 3.6 kcal mol^{-1} M^{-1} and T= 20 °C. **B.** Mechanism where dissociated hemin can bind to the unfolded state U (Appendix A.2). The values from A were used and additionally: K_{UH}^0 = 1.0 x10^{-8} M and m_{UH} = 1.7 kcal mol^{-1} M^{-1}. **C.** Mechanism where dissociated hemin self-associates to form a dimer (Appendix A.3). The values from A were used and additionally: $K_{H_2}^0$= 4.5 x10^6 M^{-1} (de Villiers et al., 2007), m_{H_2}= 1.0 kcal mol^{-1} M^{-1} (unpublished data).

3.2. Model with monodisperse free hemin

The simplest mechanism for the unfolding of a monomeric holo-globin assumes that the released hemin remains monodisperse and that, after hemin dissociation from the holoHb, unfolding of the resultant native apoHb (N) is a simple two-state process, N-to-U (Fig. 3.1A). In this case, the equilibrium constants are defined as:

$$K_{NU} = \frac{[U]}{[N]} = K_{NU}^{0} \exp\left(\frac{m_{NU}[GuHCl]}{RT}\right) \tag{3.1}$$

$$K_{NH} = \frac{[N][H]}{[NH]} = K_{NH}^{0} \exp\left(\frac{m_{NH}[GuHCl]}{RT}\right) \tag{3.2}$$

The concentration of denaturant [GuHCl] exponentially increases the K_{NH} and K_{NU} constants as defined by m_{NU} and m_{NH}, respectively. The total protein concentration at any titration point is given by:

$$P_0 = [NH] + [N] + [U] = [NH]\left(1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \tag{3.3}$$

K_{NH} is the equilibrium constant for the dissociation of hemin from native state, K_{NU} is the unfolding constant for the N to U transition, and NH, N, U, and H represent monomeric holoHb, the folded native apo-state, the unfolded apo-state, and free hemin, respectively.

Derivations of the fractions of each protein species, NH, N, U, and free H are given in the Appendix A.1. The expressions are complex because hemin binding is a bimolecular process that is favored at high initial holoHb concentrations. As a result, the free concentration of hemin, H, is a function of denaturant and total protein concentration P_0, and is defined by the following expression:

$$[H] = \frac{-(K_{NH} + K_{NH}K_{NU}) + \sqrt{(K_{NH} + K_{NH}K_{NU})^2 + 4P_0(K_{NH} + K_{NH}K_{NU})}}{2} \quad (3.4)$$

The fraction of native holo-globin is defined by:

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}} \quad (3.5)$$

where [H] is given by Equation 3.4. The fractions of N and U can be retrieved in a similar manner. As shown in Fig. 3.1A, this simple mechanism predicts that the holoHb unfolding curves should be strongly dependent on the total protein concentration. A marked right-shift of the GuHCl-induced unfolding curves will occur with increasing protein concentration if hemin remains mono-disperse after dissociation from the globin.

3.3. Model with hemin binding to the unfolded state

Expansion of the previous mechanism allows hemin to stay bound after unfolding or re-association with the unfolded state to produce a fourth UH state (Fig. 3.1B). Binding of hemin to the unfolded U state of Mb was proposed by Hargrove and Olson (1996) because they observed little protein concentration dependence for GuHCl-induced unfolding of holoMb, and this phenomenon was also observed for the thermal unfolding of cytochrome b_{562} (Robinson et al., 1998). Hemin is relatively insoluble, binds non-specifically to a variety of proteins, and is found in globin precipitates in red cells (*i.e.*, Heinz bodies) under pathological conditions (Bunn and Forget, 1986).

Hemin binding to unfolded apoHb adds another protein species and equilibrium equation to the original mechanism, and this binding and the total protein concentration are defined by:

$$K_{UH} = \frac{[U][H]}{[UH]} = K_{UH}^{0} \exp\left(\frac{m_{UH}[GuHCl]}{RT}\right) \tag{3.6}$$

$$P_0 = [NH] + [N] + [U] + [UH] = [NH]\left(1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]} + \frac{K_{NH}K_{NU}}{K_{UH}}\right) \tag{3.7}$$

The complete derivation for this four state mechanism is given in the Appendix A.2, and again a quadratic equation in [H] is generated with an even more complex root:

$$[H] = \frac{-\left(K_{NH} + K_{NH}K_{NU}\right) + \sqrt{\left(K_{NH} + K_{NH}K_{NU}\right)^2 + 4P_0\left(1 + \frac{K_{NH}K_{NU}}{K_{UH}}\right)\left(K_{NH} + K_{NH}K_{NU}\right)}}{2\left(1 + \frac{K_{NH}K_{NU}}{K_{UH}}\right)} \tag{3.8}$$

The fraction of folded holo-NH is defined as before but contains the additional [UH] term in the denominator:

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]} + \frac{K_{NH}K_{NU}}{K_{UH}}} \tag{3.9}$$

Allowing heme binding to the unfolded state decreases the stabilizing effect of increasing protein concentration because as P_0 increases, the fractions of NH and UH both increase. As a result, the unfolding curves show much less dependence on globin concentration (Fig. 3.1B). At very high protein concentrations, where P_0 ~100 times the K_d for hemin binding to the U state, there is little or no free hemin during unfolding and

the unfolding process becomes effectively unimolecular, NH to UH. Under these conditions, there is very little protein concentration dependence and the process appears to be simple two-state process (see dashed and solid lines in Fig. 3.1B).

3.4. Model with hemin dimerization

Another potential complication is that free hemin can forms dimers and higher order aggregates with itself at neutral pH, even at micromolar concentrations (Adams, 1977; Brown et al., 1970; de Villiers et al., 2007). The third mechanism in Fig. 3.1 allows self-association of free hemin to form dimers after dissociation from the holo-protein (Fig. 3.1C). A complete derivation is given in the Appendix A.3. In this case, the additional step is described by a hemin dimerization constant, K_{H_2}, which is defined by:

$$K_{H_2} = \frac{[H_2]}{[H]^2} = K^0_{H_2} \exp\left(\frac{m_{H_2}[GuHCl]}{RT}\right) \tag{3.10}$$

In this mechanism, the total free hemin also includes dimers (*i.e.*, H_{free} = [H] + 2[H_2] = [N] + [U]). As a result of this bimolecular hemin association step, the final equation for [H] is a cubic equation, and Cardano's method was used to find the appropriate root of the equation $a[H]^3+b[H]^2+c[H]+d$, which is:

$$[H] = -\frac{b}{3a} - \frac{1}{3a}\left(\sqrt[3]{\frac{2b^2 - 9abc + 27a^2d + \sqrt{\left(2b^2 - 9abc + 27a^2d\right)^2 - 4\left(b^2 - 3ac\right)^3}}{2}}\right) - \frac{1}{3a}\left(\sqrt[3]{\frac{2b^2 - 9abc + 27a^2d - \sqrt{\left(2b^2 - 9abc + 27a^2d\right)^2 - 4\left(b^2 - 3ac\right)^3}}{2}}\right) \tag{3.11}$$

using the following coefficients for the cubic equations:

$$a = 2K_{H_2}$$

$$b = 1 + 2K_{H_2}\left(K_{NH} + K_{NH}K_{NU}\right)$$

$$c = K_{NH} + K_{NH}K_{NU}$$

$$d = -P_0\left(K_{NH} + K_{NH}K_{NU}\right) \tag{3.12}$$

This value of [H] can be used to compute the fraction of native holo-NH as defined in the original three-state mechanism, Equation 3.5.

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}} \tag{3.13}$$

In general, the computed unfolding curves for this mechanism show less dependence on protein concentration compared to the mechanism where dissociated hemin remains free and monomeric in solution. The favorable effect of increasing the initial amount of holoHb on resistance to hemin loss is compensated by the increased tendency of free hemin to aggregate at higher concentrations. However, the exact dependence and shapes of the curves will be strongly dependent on how the denaturant or

condition affects the hemin dimerization reaction (*i.e.*, positive versus negative m_{H_2} values). Thus, any thorough analysis requires examining independently the dependence of K_{H_2} on pH, temperature, or denaturant.

3.5. Conclusions

The key conclusion from the theoretical curves shown in Fig. 3.1 is that hologlobin unfolding ultimately needs to be measured and analyzed carefully as function of total protein concentration. This variation is the best tool for discriminating between the three possible mechanisms. The 1 to 400 μM holo-globin concentrations were chosen to cover the normal range used in *in vitro* studies with Hbs and Mbs. In cases where unfolding curves for holoMb were measured as function of P_0, the observed dependence was small and similar to that shown in panels B and C in Fig. 3.1, indicating that dissociated hemin may bind to the unfold state (Hargrove and Olson, 1996), and/or dimerize (Asher et al., 2009; de Villiers et al., 2007), but is almost certainly never free and monodisperse in solution.

The lack of protein concentration dependence on unfolding curves of hologlobins therefore strongly suggests that heme either interacts with itself or with the unfolded protein. The more complex schemes in Figs. 3.1B-C need to be the starting points for any analysis of monomeric hologlobin unfolding, despite the apparent simplicity of the individual unfolding curves at a fixed protein concentration. These conclusions were used in the following Chapters 4 and 5.

Chapter 4

Six-state analysis for the unfolding of myoglobin

This chapter is adapted from:

Culbertson, D.S., and Olson, J.S. (2010). Role of heme in the unfolding and assembly of myoglobin. Biochemistry *49*, 6052-6063.

4.1. Introduction

Sperm whale Mb has been used extensively in structure/function studies as the classic monomeric globin. In the apoform, Mb populates a well defined intermediate in the folding pathway (See Chapter 1.3); however, the mechanism of holoMb unfolding has only been addressed in a few studies, and until our work, most analyses were semi-empirical and assumed an apparent two-state equilibrium and/or kinetic mechanism (Moczygemba et al., 2000; Ramos et al., 1999; Regis et al., 2005; Wittung-Stafshede et al., 1998). By using a set of five carefully chosen sperm whale metMb variants (*i.e.*, WT, H64F, T67P, V68T and H97D, Fig. 4.1), we have been able to demonstrate and then characterize heme binding to the apoMb unfolding intermediate. The mutations were selected to stabilize the N apoMb state but weaken hemin affinity (H64F), to selectively destabilize the N state without affecting hemin affinity or the stability of the I state (T67P), to destabilize the N state but enhance hemin affinity (V68T), and to decrease hemin affinity without affecting the stabilities of the N and I apo states (H97D).

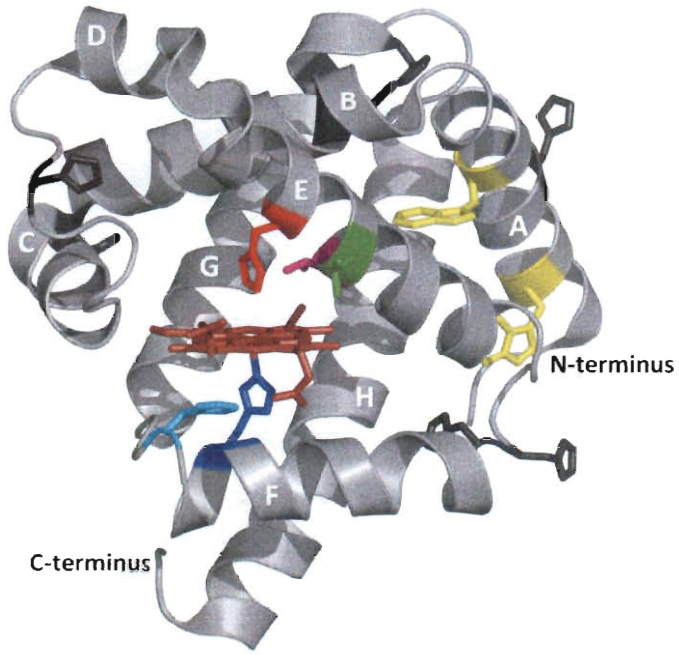


Figure 4.1. Crystal structure of sw aquometMb. Ribbon representation with residues of interest indicated in sticks: His64 (red), Thr67 (green), Val68 (purple), His97 (cyan), His93 (navy blue), Trp7 and Trp14 (yellow). The other His residues are indicated in black sticks. The A-H helices are labeled in white. Heme is represented in brown sticks. Image created using PyMol and PDB entry: 1jp6.

Analysis of GuHCl-induced unfolding of these variants requires a six-state mechanism containing native (N), intermediate (I), and unfolded (U) apoMb states and their hemin bound counterparts, NH (ferric holoMb), IH, and UH. This six-state mechanism applies to all monomeric proteins containing non-covalently bound heme with an intermediate regardless of the exact nature of the cofactor binding, provides a framework for quantifying the factors that contribute to the overall stability of the holoprotein, and allows independent and accurate determination of equilibrium constants for hemin binding to native and intermediate apoprotein states.

4.2. Three-state analysis of apoMb unfolding in GuHCl

As described in Chapter 1, apoMb populates an intermediate in the equilibrium-unfolding pathway. The unfolding process of apoMb can be followed by CD (secondary structure content) and fluorescence emission (environment of Trp7 and Trp14, see Fig.

4.1). We have assumed that the signals from both detections report on the same intermediate, and, as described by Ramsay *et al.* (1995), Scott *et al.* (2000), and Smith (2003), the unfolding data can be fit to a two-step/three-state N-to-I-to-U equilibrium-unfolding mechanism analogous to the one originally developed by Barrick, Baldwin, and colleagues (Barrick and Baldwin, 1993a; Barrick et al., 1994; Griko et al., 1988; Nishii et al., 1995).

The amount of helical content was analyzed by measuring ellipticity at 222 nm. Changes in the environment of the Trp side chains were followed by measuring fluorescence emission at 341 nm, which gives rise to a bell-shaped curve with an increase in fluorescence for the intermediate I state and a decrease for the completely unfolded or U state (see Fig. 4.2; Hargrove et al., 1994a; Ramsay et al., 1995; Scott et al., 2000). This rise and fall in fluorescence intensity allows for a better definition of the three-state unfolding parameters. The origin of hyperfluorescence from the intermediate state still remains unclear. Fluorescence of WT apoMb has been suggested to reflect emission from Trp7 centered at 333 nm, which is quenched by Lys79 and/or Met131 in the folded state but not in the intermediate, and emission centered at 321 nm from Trp14, which is buried in the apolar region between the A, G, and H helices in both the N and I states and only quenched after unfolding (Hargrove et al., 1994a; Irace et al., 1981; Kirby and Steiner, 1970; Tcherkasskaya et al., 2000a; Tcherkasskaya et al., 2000b). Alternatively, the hyperfluorescence of the intermediate could be caused by enhanced flexibility of the fluorophores in the intermediate state (Ervin et al., 2002).

The stability of the apoMb variants are defined by the K_{NI}^{0} and K_{IU}^{0} equilibrium stability constants for the N-to-I and I-to-U transitions, respectively, and the linear

dependences of the corresponding free energies on GuHCl concentrations are defined by the m_{NI} and m_{IU} values.

$$K_{NI} = \frac{[I]}{[N]} = K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) \tag{4.1}$$

$$K_{IU} = \frac{[U]}{[I]} = K_{IU}^{0} \exp\left(\frac{m_{IU}[GuHCl]}{RT}\right) \tag{4.2}$$

The observed signal at a given [GuHCl] is the weighted sum of the observable signals, S_N, S_I, and S_U, for each state (Barrick and Baldwin, 1993b; Hargrove et al., 1994a; Scott et al., 2000; Smith, 2003), which is expressed as:

$$S = \frac{S_N + S_I K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) + S_U K_{NI}^{0} K_{IU}^{0} \exp\left(\frac{(m_{NI}+m_{IU})[GuHCl]}{RT}\right)}{1 + K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) + K_{NI}^{0} K_{IU}^{0} \exp\left(\frac{(m_{NI}+m_{IU})[GuHCl]}{RT}\right)} \tag{4.3}$$

The m_{NI} and m_{IU} values for the N-to-I and I-to-U transitions in apoMb were fixed to 2.35 and 1.36 kcal mol^{-1} M^{-1}, respectively (Scott et al., 2000; Smith, 2003), assuming that the single point mutations do not significantly affect the change in hydrophobic exposure between the N, I, and U states. Only the unfolding parameters, K_{NI}^{0} and K_{IU}^{0}, were allowed to modulate the N-to-I, and I-to-U transitions of the apo-variants.

Definition of the signal parameters for analysis of apoMb unfolding curves using the three-state mechanism

The CD and fluorescence curves for unfolding of the five apoMb variants were fit to the three-state N-to-I-to-U mechanism using Equation 4.3 and in Fig. 4.2. The fitted signal parameters for the N, I, and U states are listed in Table 4.1 and obtained from the

global fits of the CD data at 222 nm (CD) and the fluorescence data at 341 nm (F). The unfolding equilibrium constants are given in Table 4.1. The fluorescence emission of the molten globule I state intermediate at this wavelength, $S_I(F)$, varies significantly and is ~1.8 to 3.6-fold higher than that of unfolded state, $S_U(F)$ which was normalized to be 1.0 for all the variants.

Table 4.1. Respective circular dichroism and fluorescence signals for the three states of the apoMb variants

ApoMb variant	WT	H64F	T67P	V68T	H97D
$S_N(CD)^1$	1.00	1.22	0.75	1.08	1.02
$S_I(CD)^1$	0.42	0.78	0.57	0.34	0.41
$S_U(CD)^1$	0.00	0.00	0.00	0.00	0.00
$S_N(F)^1$	1.03	1.09	0.55	1.17	0.92
$S_I(F)^1$	1.97	1.79	2.18	2.17	3.57
$S_U(F)^1$	1.00	1.00	1.00	1.00	1.00

[1]The CD_{222nm} value of wild-type apoMb N state was normalized to 1.0 (*i.e.*, $S_N(CD)$ WT =1.0). The $S_N(F)$ and $S_I(F)$ values were allowed to vary, but all the $S_U(F)$ values were set to 1 for emission at 341 nm. The $S_I(CD)$ values were allowed to vary but in some cases are poorly defined when the unfolding transition appears concerted; however, all the $S_I(F)$ values are well defined due to the bell-shape the fluorescence curves.

4.3. Equilibrium-unfolding of apoMb variants

Overlays of both detections for all apo-variants are shown in Fig. 4.2. As expected, the unfolding data for the five apoMb variants were readily interpreted in terms of a three-state model. In most cases, the CD data reveal an inflection point in the unfolding curves indicative of a populated intermediate; however, the CD data alone are not sufficient to define accurately the stability parameters for the N-to-I, and I-to-U transitions. The fluorescence unfolding curves demonstrate unequivocally the presence of an intermediate, which is characterized by hyperfluorescence centered at a wavelength in

between the fluorescence peaks for the native N and unfolded U states, as was observed previously (Hargrove et al., 1994a; Ramsay et al., 1995; Scott et al., 2000).

We successfully obtained fitted K_{NI}^{0} and K_{IU}^{0} stability constants for the N-to-I, and I-to-U transitions in the absence of denaturant, and the values are listed along with free energy parameters in Table 4.1. The shape of the unfolding curve is directly influenced by the population fractions and respective signals for each of the three states (Equation 4.3). The populations of each state are markedly dependent on the stability constants K_{NI}^{0} and K_{IU}^{0}, which vary between the mutants, whereas the *m*-values for the two transitions are relatively invariant and were kept constant. The respective signals for the folded (N) and intermediate (I) states of each apoprotein vary due to the structural changes induced by the point mutation. These effects are most noticeable for the hyperfluorescence intensity of the molten globule intermediate in H64F, which is significantly less intense than the one in H97D. This decrease in intermediate fluorescence for H64F could reflect a more compact molten globule due to its increased core hydrophobicity and stability. The CD unfolding curves also reveal a disparity between the higher helical content of the N and I states of H64F when compared to lower helical content of for the same T67P states (See Table 4.1).

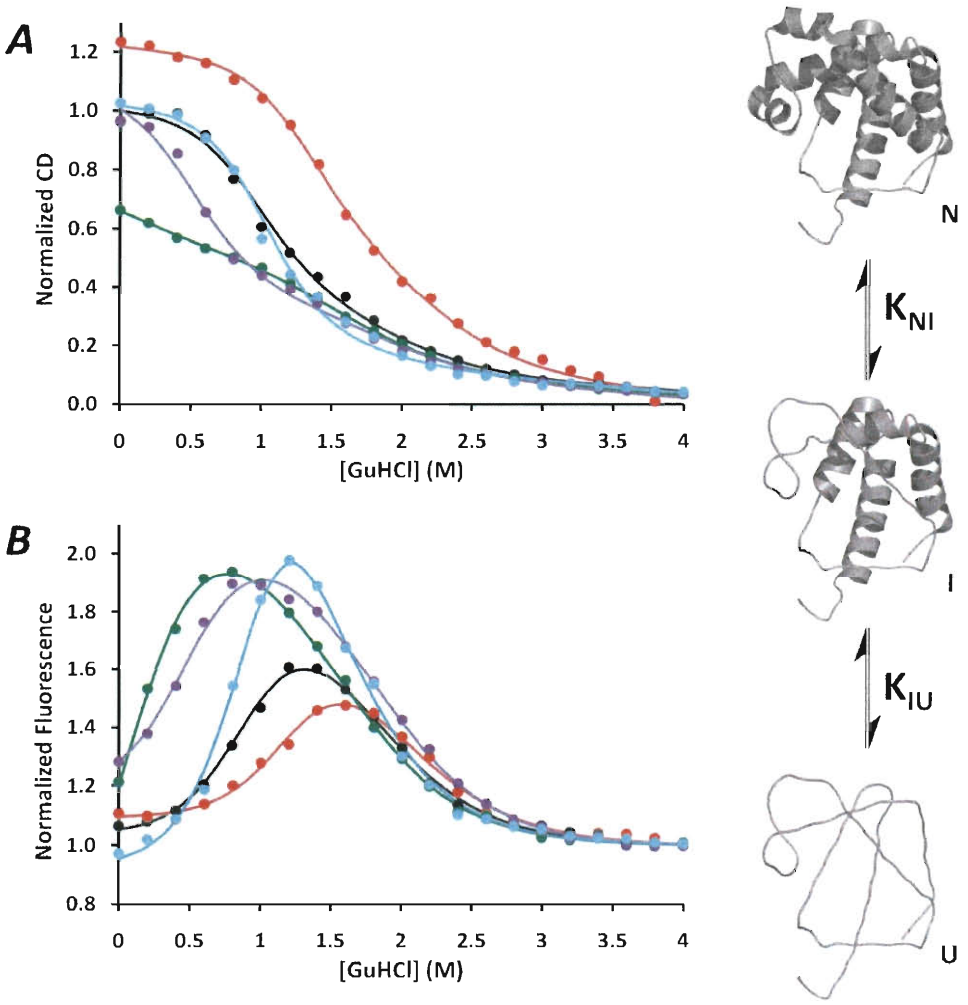


Figure 4.2. GuHCl-induced equilibrium-unfolding of apoMb variants. The unfolding curves obtained by normalizing the changes in Far-UV CD at 222 nm to 1.0 for native WT apoMb and assuming that the U states of all five variants have the same CD_{222nm} (left panel) and by normalizing fluorescence emission intensities at 341 nm to 1.0 for the final, completely unfolded state of each mutant (right panel). Global fits to a three-state mechanism are shown for WT (black), H64F (red), T67P (green), V68T (purple) and H97D (cyan). Conditions: 10 μM protein, 10 mM potassium phosphate buffer pH 7, 20 °C. The three-state unfolding mechanism is depicted on the right side, with the structures of N, I, and U states. The structures were created using PyMol and the PDB entry: 1jp6 for WT metMb, taking into account the secondary structure present in the N state (Cocco and Lecomte, 1990; Eliezer and Wright, 1996; Hughson et al., 1990) and in the I state (Hughson et al., 1990; Jennings and Wright, 1993). The thick ribbons indicate intact helical secondary structure and the thin lines represent unfolded structures. The U state is considered completely unfolded.

Table 4.2. Stability parameters for the apoMb variants[1]

ApoMb variant	K_{NI}^0	K_{IU}^0	$1/K_{NU}^0$
WT	0.021	0.019	2500
H64F	0.0061	0.011	15000
T67P	0.67	0.029	51
V68T	0.13	0.017	450
H97D	0.013	0.061	1300

[1] The GuHCl-induced unfolding measurements were carried out in 10 mM potassium phosphate pH 7, 20 °C. The K_{NI}^0 and K_{IU}^0 were obtained by fitting simultaneously both CD and Fluorescence signals for the unfolding of the apoMb variants to the three-state unfolding mechanism using Equation 4.3 (See Appendix A.4). The values of m_{NI} and m_{IU} were set to 2.35 and 1.36 kcal mol^{-1} M^{-1}, respectively.

The variants show a wide range of stabilities for the N states, with the H64F mutant being the most stable and the T67P variant being the least stable. As expected, K_{IU}^0 is much less affected by these distal pocket mutations than K_{NI}^0, supporting the view that the heme pocket is melted in the molten globule intermediate I state. The H64F substitution increases the hydrophobicity of the heme pocket, inhibiting its unfolding and decreasing K_{NI}^0 3-fold (Table 4.2). The H97D variant displays stabilities for the N and I states that are similar to those of WT apoMb because the imidazole side-chain of residue 97 is solvent exposed. The V68T substitution increases the polarity of the heme pocket by inserting an additional polar side chain adjacent to the ligand binding site, and causes a marked decrease in the stability of the N state, which is reflected by the 6-fold increase in K_{NI}^0 (Hargrove et al., 1994a). Finally, the T67P variant introduces a "kink" in the E helix lowering its stability and that of the entire heme pocket in the N state, resulting in a dramatic 30-fold increase in K_{NI}^0.

4.4. Analysis of holoMb equilibrium-unfolding data using a six-state mechanism

The general scheme for the unfolding of a monomeric heme protein is shown in Fig. 4.3, and allows heme, H, to bind to all three apoprotein states: N, I, and U. In this equilibrium model, the K_{NI}^{0} and K_{IU}^{0} stability constants and m_{NI} and m_{IU} values for the N-to-I, and I-to-U transitions of apoMb were determined independently in separate experiments with the apo-forms and not allowed to vary (Fig. 4.2).

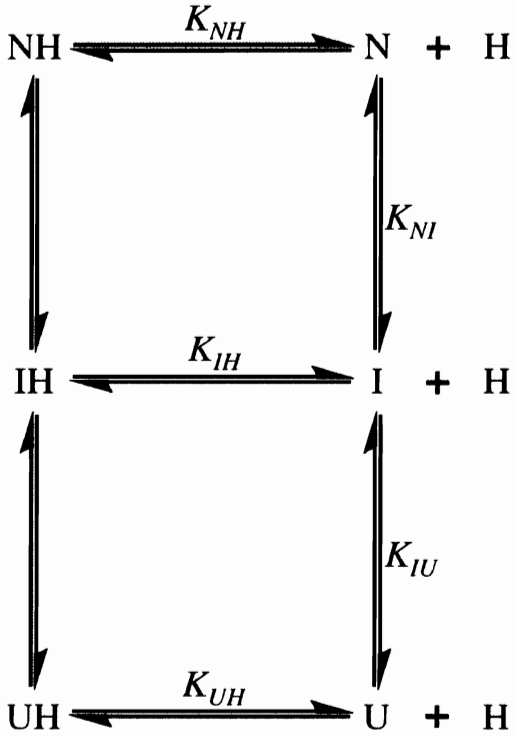


Figure 4.3. Six-state mechanism for the analysis of holoMb equilibrium unfolding. The scheme is based upon the three-state unfolding mechanism of apoMb and the assumption that heme can bind to all three apoMb states.

Unfolding curves for the holoMb variants were analyzed according to the six-state mechanism. The heme dissociation constants from all possible heme-bound states, NH (metMb), IH, and UH were allowed to vary in order to define the equilibrium unfolding curves for holoMb. The three hemin dissociation constants are defined as:

$$K_{NH} = \frac{[N][H]}{[NH]} = K_{NH}^{0} \exp\left(\frac{m_{NH}[GuHCl]}{RT}\right) \tag{4.4}$$

$$K_{IH} = \frac{[I][H]}{[IH]} = K_{IH}^{0} \exp\left(\frac{m_{IH}[GuHCl]}{RT}\right) \tag{4.5}$$

$$K_{UH} = \frac{[U][H]}{[UH]} = K_{UH}^{0} \exp\left(\frac{m_{UH}[GuHCl]}{RT}\right) \tag{4.6}$$

The dependences on GuHCl concentrations are represented by m_{NH}, m_{IH}, and m_{UH}. To our knowledge, *m*-values for hemin dissociation from globins have never been reported, and our determination of these parameters is discussed later in this chapter.

The population fractions of each species were computed by first obtaining the free concentration of heme [H] by solving for the roots of a complex quadratic equation, which leads to the following expression:

$$[H] = \frac{-\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)}{2\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)} + \frac{\sqrt{\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)^2 + 4P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)}}{2\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)} \tag{4.7}$$

The fraction of each species was calculated using the free heme concentration [H], the two apoMb stability constants, and the three heme equilibrium dissociation constants:

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NI}}{[H]} + \frac{K_{NH}K_{NI}K_{IU}}{[H]} + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}} \tag{4.8}$$

The fraction of free hemin (Y_H) is given by the sum of the fractions of apoglobin states (Y_N+Y_I+Y_U). Each of the stability and affinity constants depicted in Equations 4.7-8 were computed at different [GuHCl] using the *m*-values: m_{NI}, m_{IU}, m_{NH}, m_{IH}, and m_{UH} as defined in Equations 4.1-2 and 4.4-6 to determine the fractions of each species as a function of [GuHCl].

The equilibrium-unfolding curves for the ferric holoMb variants were also obtained at single wavelengths. The amount of helical content was analyzed using ellipticity at 222 nm. For fluorescence, large emission increases observed at 355 nm reflect the loss of hemin, which, when bound in the native folded state, quenches almost all Trp fluorescence. For absorbance detection, the unfolding data were analyzed at the Soret wavelength maximum of the ferric state for each variant: 409 nm for WT, T67P, V68T and H97D, and 393 nm for H64F pentacoordinate metMb. The three detections were fit simultaneously according to their respective computed signals from:

$$S = S_{NH}Y_{NH} + S_NY_N + S_{IH}Y_{IH} + S_IY_I + S_{UH}Y_{UH} + S_UY_U + S_HY_H \qquad (4.9)$$

Fitting of the holoMb variants to the six-state mechanism.

The circular dichroism (CD), fluorescence (F) and Soret absorbance (Abs) signals for the unfolding of the holoMb variants were fit to the six-state mechanism using Equations 4.1-2 and 4.4-7. The signals from all three detections were simultaneously fit using Equation 4.9, and the respective signals for each state were either fixed to values obtained from the analysis of apoMb unfolding or allowed to vary. The signal parameter values used to generate the fitted curves in Fig 4.2 are listed in Table 4.3 and the equilibrium constants for hemin dissociation and *m*-values are given in Table 4.4.

The S(CD) values are different from those used in the analysis of apoMb. In this case, the CD_{222nm} of WT apoMb was normalized to 1.0 and is virtually the same for the other four variants. The $S_N(CD)$ values for the native apoMb states shown in Table 4.1 were renormalized relative to that for WT holoMb and are in range of 0.47 to 0.71, *i.e.*, roughly 60% of the CD of the holoprotein. Similarly, the re-normalized $S_I(CD)$ values are smaller, ~ 20-30% of the holoprotein, but still roughly half of that for the apo-N states. The CD signal of the IH state, $S_{IH}(CD)$, was allowed to vary but remained in the range of that for the apo-I state. The CD signals for the UH state were set equal to zero because the population of this species is never significant at the holoMb concentration used (P_0=10 μM), whereas $S_U(CD)$ for the U state was allowed to vary starting from 0.0. As shown both $S_U(CD)$ and $S_{UH}(CD)$ are effectively zero.

For holoMb unfolding, the fluorescence at 355 nm of the apo-U state was normalized to 1.0 for each variant. Thus, all the $S_U(F)$ values were set to 1.0 and allowed to vary but never deviated from this value. The fluorescence signals of the holoMb states, $S_{NH}(F)$, were initially set to 0.0, allowed to vary, but never deviated significantly from zero. The IH state signal, $S_{IH}(F)$ was found to be higher than expected (*i.e.*, 0.36 to 0.69), considering the fact that heme quenches virtually all of the fluorescence of the native holoMb NH state. The fluorescence signals of the apo-N and apo-I states were initially estimated from analyses of apoMb unfolding fluorescence data measured at 355 nm instead of at 341 nm as was done in Table 4.1. The fluorescence of the intermediate state is less prominent at 355 nm than at 341 nm. The fitted values of $S_N(F)$ are roughly 50% of the values obtained from the analysis of apoMb unfolding data due to the change in emission wavelength, and the $S_I(F)$ values are also smaller for the same reason. The only

unexpected results are the large values of $S_{IH}(F)$, which are ~50% of the value of the U state. As described in the Main Text, we expected that the presence of hemin bound to the I state would still result in almost complete quenching. However, that is not the case for the intermediate hemichrome species.

Table 4.3. Respective far-UV CD, Fluorescence emission, and Soret Absorbance Signals used in the fits of the holoMb variants to the six-state unfolding mechanism.

HoloMb variant	WT	H64F	T67P	V68T	H97D
$S_{NH}(CD)$	1.00	0.98	1.01	0.99	1.03
$S_N(CD)^1$	0.60	0.71	0.47	0.63	0.61
$S_I(CD)^1$	0.25	0.46	0.36	0.20	0.25
$S_{IH}(CD)$	0.55	0.22	0.25	0.32	0.17
$S_U(CD)^1$	0.03	0.03	0.02	0.02	0.02
$S_{NH}(F)^2$	0.01	0.04	0.00	0.02	0.02
$S_N(F)$	0.61	0.80	0.82	0.50	0.64
$S_I(F)$	1.24	1.20	1.46	1.10	1.56
$S_U(F)^2$	0.97	0.96	0.97	0.98	0.99
$S_{IH}(F)$	0.36	0.69	0.46	0.53	0.66
$S_{NH}(Abs)^3$	1.00	1.00	1.02	1.00	1.01
$S_{IH}(Abs)$	0.69	0.17	0.20	0.35	0.15
$S_H(Abs)$	0.03	0.02	0.01	0.01	0.01

[1] The signals for $S_N(CD)$, $S_I(CD)$, $S_U(CD)$ were re-normalized from Table 4.1 assuming $S_{NH}(CD)$ is 1.0 for WT holoMb and $S_N(CD)$ is 0.60 for WT apoMb as described in the text above.

[2] The fluorescence signal for the U states were defined as 1.0 experimentally, allowed to vary, and did not deviate from this value. The fluorescence signal from the NH state was set to 0.0, allowed to vary, and did not change from this value.

[3] $S_N(Abs)$, $S_I(Abs)$, and $S_U(Abs)$ were set to 0.0 because no hemin is bound. The WT $S_{NH}(Abs)$ value was initially set to 1.0 (normalized total absorbance change). The other parameters were initially defined experimentally and then allowed to vary as described in the text above.

The Soret absorbance change for each variant was normalized to 1.0 at the wavelength of observation. Thus, the signal for the NH state, S_{NH}(Abs), was defined as 1.0. The signal for free H was set to zero initially, allowed to vary, but, as expected, remained effectively at zero since the final state at high [GuHCl] is primarily U+H. The S(Abs) values for all the apoMb species are defined and fixed at 0.0. The initial value of the hemichrome IH state, S_{IH}(Abs) was set to a value estimated from the model hemichrome spectrum shown in Fig. 4.6C and the initial and final spectra shown in panels A and B of Fig. 4.6, and then allowed to vary. In general, the fitted values are very similar to the initial starting values. The Soret absorbance signal of the UH state, S_{UH}(Abs), was initially set to 0.2, allowed to vary, but did not change. There is never any significant amount of the UH state even at high [GuHCl], making S_{UH}(Abs) undefined.

The six-state mechanism is required to compare holoMb variants

In Chapter 3, we used a four-state mechanism to analyze the equilibrium-unfolding of holoMb variants, whereas in this Chapter, we have adopted a six-state mechanism because the unfolding of apoMb clearly shows the presence of an intermediate, even for wild-type protein. To test the need for a more complex model, I fit the unfolding curves for the WT and H97D holoMb variants to the simpler four-state mechanism (as in Fig. 3.1B). First the apoMb CD unfolding data were fit to a two-state unfolding mechanism (not shown) to obtain values for K_{NU}^{0} and m_{NU}, which were: K_{NU}^{0} = 0.091 and 0.067, and m_{NU} = 1.2 and 1.5 kcal mol^{-1} M^{-1}, respectively, for WT and H97D apoMb variants. These fits were poor and could not reproduce the "bell-shaped" fluorescence data but were necessary to estimate apoMb unfolding parameters for analysis of holoMb data.

As expected, the unfolding of WT holoMb can be fit reasonably well to a four-state unfolding mechanism (Fig. 4.4, upper panels). The distribution of the residuals (observed minus calculated signals) indicates that both models can represent the WT data, however, as described above, the lack of an intermediate in the unfolding mechanism for the apo form argues against using this simpler model. However, the fits in the upper panels of Fig. 4.4 for WT holoMb do provide a reasonable estimate for $K_{NH}^{0} \approx 1 \times 10^{-14}$ M with $m_{NH} \approx 6.2$ kcal mol^{-1} M^{-1}.

In contrast, the four-state model cannot accurately describe the unfolding of H97D holoMb. The fits to this simpler model are poor when compared to those to the six-state (Fig. 4.4, bottom panels). The residuals for the four-state model fits are large and systematic, especially for the fluorescence data. However, even with the poor fits, the four-state analysis still indicates that the H97D mutant has a much lower affinity for hemin than WT Mb, but the absolute value of K_{NH}^{0} is several fold higher than the previously estimated value of 2.2×10^{-11} M (Hargrove and Olson, 1996).

In our view, the four-state mechanism is not appropriate for a thorough quantitative analysis of the unfolding curves of the holoMb variants because intermediate states are observed with or without bound heme. The three states that populate during unfolding of the apo forms are well defined by fluorescence and CD, and the presence of an observable hemin-bound intermediate proves that the use of a six-state mechanism is required.

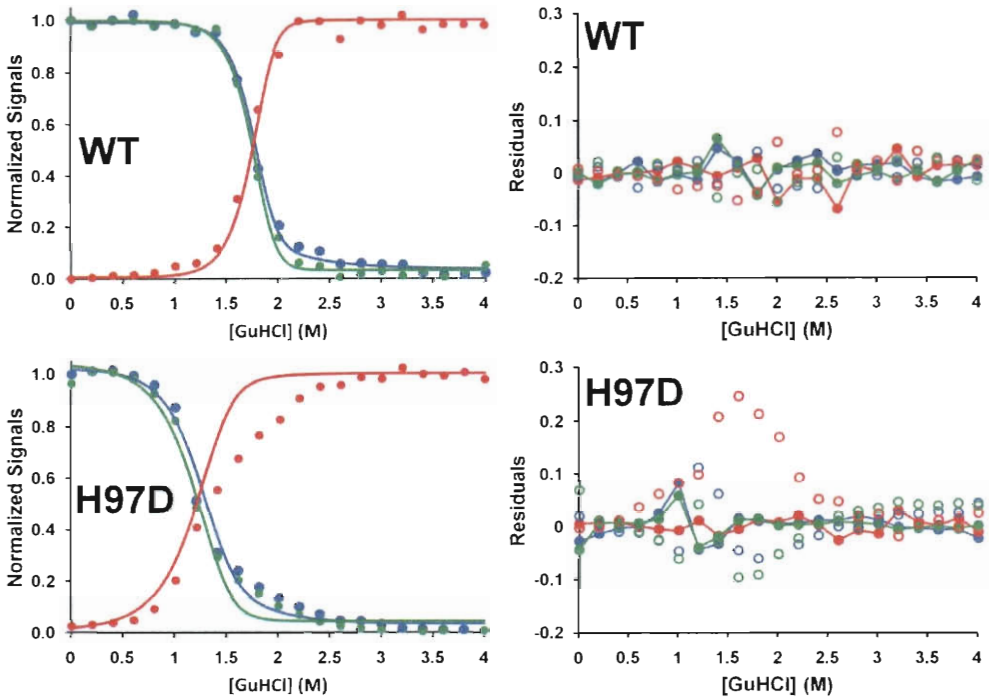


Figure 4.4. Four-state mechanism for the unfolding of WT and H97D holoMb variants. **LEFT:** Unfolding curves by CD (blue), fluorescence (red), and Soret absorbance (green) for WT and H97D variants, as shown in Fig. 4.5. **RIGHT:** Respective fitting residuals to the 4-state mechanism (open circles) and to the 6-state mechanism (filled circles connected with lines), using the same colors as in the left panels.

Protein concentration

In general, holoMb unfolding should depend on total protein concentration because it involves dissociation of heme, which should be facilitated at low concentrations. For this chapter, we focused on resolving whether hemin may bind to the intermediate state using the relative differences of apoglobin stabilities and hemin affinities between the variants, and thus, we neglected the added complexities of hemin dimerization and used a single protein concentration P_0= 10 μM. The addition of a free hemin dimerization step after unfolding greatly complicates the analysis, and is described in detail in Chapter 5.

4.5. Unfolding of WT holoMb

The agreement between the observed data and the fitted curves is remarkably good (Fig. 4.4), considering that the key apoMb unfolding parameters were fixed to the values listed in Table 4.2. The fractions of each state during unfolding are also shown in Fig. 4.4. The complete set of data for all five variants demonstrates unequivocally that unfolding of holoMb is not a one-step process and involves equilibrium populations of IH and I states at moderate denaturant concentrations. However, as observed previously, the curves for WT holoMb by themselves are not definitive with respect to the presence of an IH state.

The dissociation constant K_{NH}^{0} for WT holoMb was estimated to be ~0.9 x10^{-13} M at pH 7, 20 °C, which is in reasonable agreement with the equilibrium dissociation constant, $K_{.H}$ ~0.3 x10^{-13} M, computed from the ratio of association and dissociation rate constants for WT metMb at pH 7, but at higher salt concentration in 0.45 M sucrose at 37 °C (Table 4.2, Hargrove et al., 1996a; Hargrove and Olson, 1996). In addition, we were able to successfully estimate K_{IH}^{0}, which is ~1.4 x10^{-11} M, indicating that the hemin affinity of the WT I state is ~100 fold weaker than that of the N state. The value of K_{UH}^{0} is poorly defined but estimated to be ~10^{-6} M for all five variants, in agreement with previous estimates of heme binding to unfolding apoMb and apocytochrome b_{562} (Hargrove and Olson, 1996; Robinson et al., 1997).

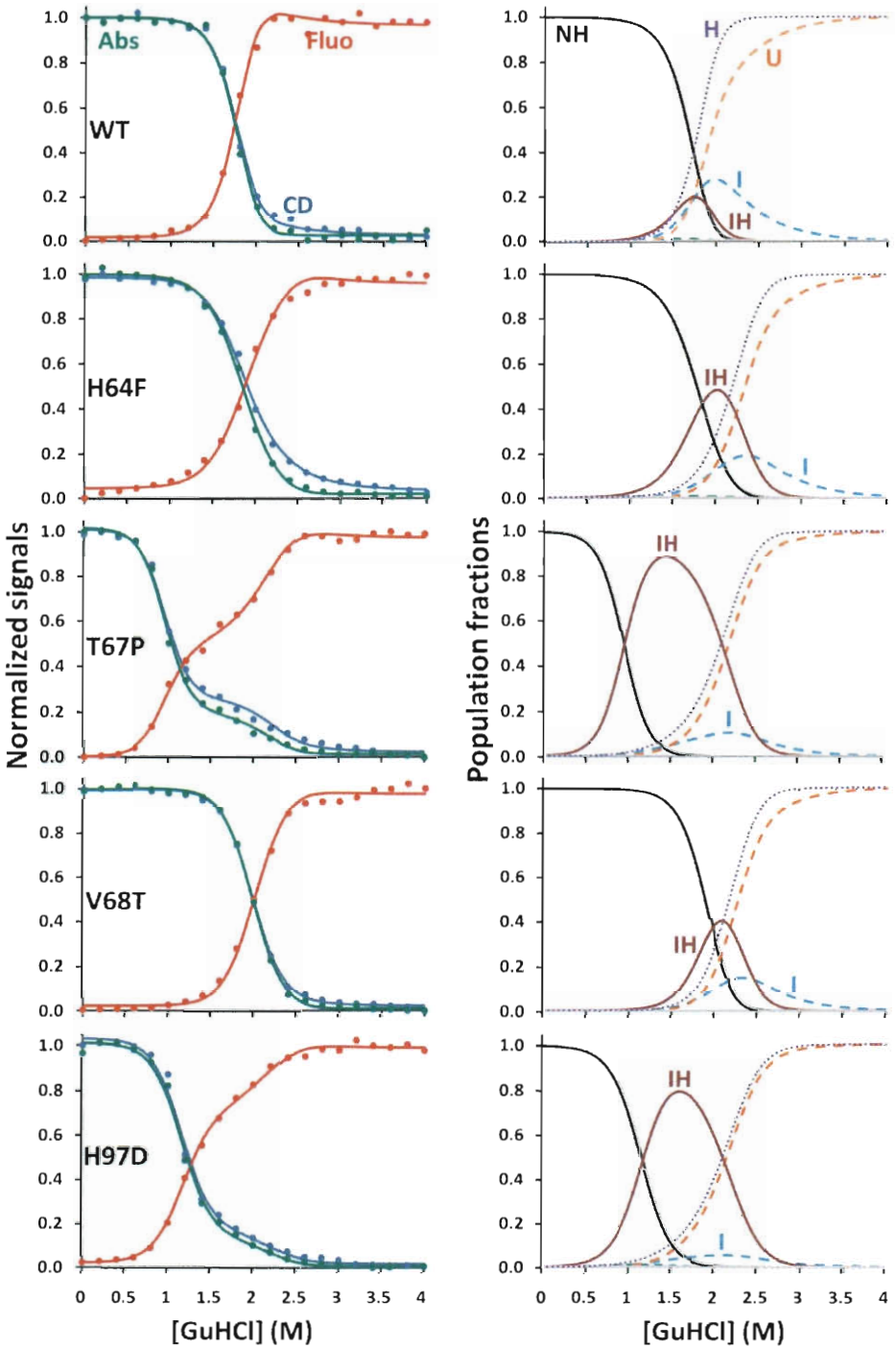


Figure 4.5. GuHCl-induced equilibrium unfolding of holoMb variants. The left panels represent the unfolding data from far-UV CD at 222 nm (blue), fluorescence emission at 355 nm (red), and Soret absorbance (green), with the best fit to the six-state unfolding mechanism. The right panels depict the population fractions of each species with regards to [GuHCl]. The conditions were: 10 μM protein, 10 mM potassium phosphate pH 7, 20 °C.

Table 4.4. Hemin dissociation constants and *m*-values for the NH and IH states of holo-metMb variants[1].

HoloMb variant	K_{NH}^0 (M)	m_{NH} kcal mol^{-1} M^{-1}	K_{IH}^0 (M)	m_{IH} kcal mol^{-1} M^{-1}	K_{-H} pH 7-8[2] (M)	K_{-H} pH 5[2] (M)
WT	0.88×10^{-13}	4.4	1.4×10^{-11}	4.0	0.28×10^{-13}	28×10^{-13}
H64F	4.1×10^{-13}	3.5	1.0×10^{-11}	3.3	2.2×10^{-13}	130×10^{-13}
T67P	1.0×10^{-13}	4.3	3.6×10^{-11}	2.8	n.d.	n.d.
V68T	0.016×10^{-13}	4.1	1.5×10^{-11}	3.2	$\leq 0.05 \times 10^{-13}$	1.1×10^{-13}
H97D	84×10^{-13}	3.4	8.8×10^{-11}	2.4	180×10^{-13}	1100×10^{-13}

[1] The unfolding measurements were carried out in 0.01 M potassium phosphate pH 7, 20 °C. K_{UH}^0 and m_{UH} were fixed to 1.0×10^{-6} M and 2.4 kcal mol^{-1} M^{-1}, respectively, for all five variants. The CD, fluorescence, and Soret absorbance signal parameters are given in Table 4.3.

[2] The values for K_{-H} were computed from the hemin dissociation rate constants assuming that the association rate constant is 1.0×10^8 M^{-1} s^{-1} for all five variants, as described in Hargrove et al. (1996a). The dissociation rate constants were taken from data at pH 7 and pH 5 at 37 °C in 0.45 M sucrose and 0.15 M sodium phosphate or acetate, respectively (Hargrove and Olson, 1996; Hargrove et al., 1996b).

4.6. H64F, a variant with increased apoMb stability but decreased hemin affinity

The unfolding of the H64F mutant is depicted in Fig. 4.5. This mutant was selected because the Phe substitution increases the stability of the N state of the apo-form by replacing the polar imidazole side chain with a relatively apolar aromatic ring which excludes water from the heme pocket (Hargrove et al., 1994a; Quillin et al., 1993). However, removal of the distal histidine eliminates the His64 hydrogen bond to the water molecule that is coordinated to the hemin iron atom. As a result, water is no longer bound to the iron atom, and H64F metMb is pentacoordinate, showing a much weaker Soret band at 393 nm instead of the strong 409 nm band seen for the aquomet form of WT metMb. In addition, the affinity of the N state of H64F apoMb for hemin is ~4-fold smaller than that of WT apoMb. In terms of dissociation equilibrium constants, K_{NH}^0(H64F) is ~4×10^{-13} M compared to K_{NH}^0(WT) is ~0.9×10^{-13} M.

As is observed for WT holoMb, the unfolding curves for H64F holoMb do not have obvious inflection points; however, the unfolding curves are clearly broadened and the transition midpoints occur at a slightly higher [GuHCl] than those for WT holoMb. Fitting to the six-state model suggests the presence of a significant population the IH state during unfolding, which causes broadening of the overall unfolding transition. Fitting to a two-state model would require a significant decrease in the overall unfolding *m*-value, which is unlikely to be caused by a single point mutation. The larger value of K_{NH}^{0} for H64F metMb is in agreement with the larger value estimated from the ratio of the dissociation and association rate constants for hemin binding to this mutant (Table 4.4, Hargrove et al., 1996b). In contrast, the equilibrium constant for hemin dissociation from the mutant IH state, K_{IH}^{0}, is estimated to be 1.0 x10^{-11} M, which is similar to that for the WT Mb intermediate (Table 4.4).

The presence of an IH state during unfolding of H64F holoMb is suggested by a red-shift of the Soret band to 410 nm at a GuHCl concentration favoring ~50% net unfolding. We deconvoluted the UV-visible absorbance spectra of H64F holoMb/GuHCl mixtures near the unfolding midpoint to obtain the absolute spectrum of the IH state. The result is shown in Fig. 4.6B. Interestingly, the deconvoluted IH spectrum is characteristic of a hemichrome species with a Soret maximum at 415 nm, a major low spin β band at 535 nm, and an α band shoulder at 565 nm (Rifkind et al., 1994). Thus, we propose that the IH state in Mb is a hemichrome. Unfolding of the heme pocket of the holoprotein appears to allow formation of a bis-histidyl linkage to the heme iron even in the absence of a distal histidine. Thus, the hemichrome-like IH state may involve bis-His coordination by alternative sets of two histidines in the heme pocket region, and not just His64 and

His93 (Fig. 4.6). Alternatively, the bis-axial coordination could arise from a Met/His pair, and there is one methionine side chain in the heme pocket region, Met55(D5).

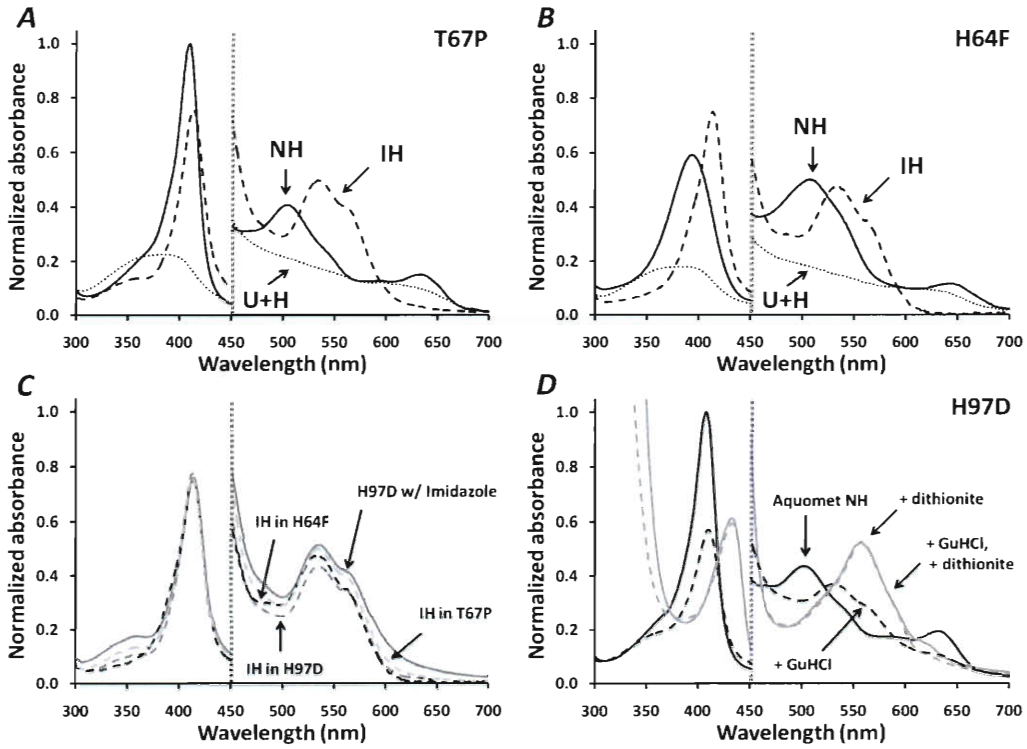


Figure 4.6. Evidence of a hemichrome species as a hemin-bound intermediate. **Panels A and B**, Absorbance spectra for T67P, H64F in the native folded NH, IH, and U + H states. All absorbance values were normalized to the peak absorbance of WT metMb at 409 nm at 10 μM total heme, which was the concentration used for all the samples. The normalized scale for the visible region, 450 to 700 nm, is 0 to 0.15. Deconvoluted spectra, for the IH states were computed assuming the starting NH and final U+H states shown in the panels and then subtracting weighted sums of these spectra from the spectra measured at the midpoint GuHCl concentration where the fraction of IH state is the largest. The fractions of NH, IH, and free H at the intermediate GuHCl concentration were estimated from the fitted parameters in Table 4.4 and Equations 4.8-9. **Panel C**, Overlay of the deconvoluted spectra of the IH states for T67P, H64F, and H97D Mb obtained from the GuHCl titrations and the spectrum of imidazole bound to H97D metMb (at 100 mM free imidazole). **Panel D**, spectra for H97D holo-metMb in the absence (solid line) and presence (dashed line) of 1.4 M GuHCl, before reduction (black) and after addition of sodium dithionite (gray). The buffer conditions were 10 mM potassium phosphate pH 7, 20 °C.

4.7. T67P, a variant with wild-type hemin affinity but an unstable N state

The unfolding of the holo T67P mutant is shown in the middle panels of Fig. 4.5. This mutant was selected as a variant with a highly destabilized apoglobin N state (Table 4.2, largest K_{NI}^0 value) due to placement of proline in the middle of the E helix (Fig. 4.1). However, the affinity of the apoprotein N state for hemin is similar to that of WT Mb. The unfolding curves of T67P holoMb all show well-defined inflection points, which demonstrate unambiguously the presence of an intermediate species. These data were readily fit to the six-state model, and the major intermediate species is the IH state, which dominates at ~1.4 M GuHCl. The IH state for this mutant is favored because of the instability of the apoMb N state, which facilitates partial unfolding even with the hemin-cofactor bound. Analysis of the UV-visible absorbance spectra during unfolding reveals directly that the IH state is a hemichrome-like species, which is characterized by distinct absorbance bands at 415, 535, and 565 nm (Rifkind et al., 1994). For this mutant, the hemichrome spectrum can be observed almost by itself at the midpoint [GuHCl] value because the fraction of IH is ~ 0.90 (Fig. 4.5). The deconvoluted spectrum of the IH state for T67P metMb is shown in Figs. 4.6A and 4.6C.

The fitted parameters for T67P metMb unfolding are listed in Tables 4.3 and 4.4. As expected, this mutation does not affect heme affinities, and the computed hemin dissociation constants remain similar to those for WT metMb with K_{NH}^0 ~1.0 x10^{-13} M and K_{IH}^0 ~3.6 x10^{-11} M (Table 4.4). Only the apoMb N-to-I unfolding parameters are affected by the T67P substitution (Table 4.2). However, the large increase in K_{NI}^0 causes a marked effect on the holoMb unfolding curves and a dramatic increase in the population of the IH

state, allowing its spectral properties to be characterized and assigned unambiguously to a hemichrome species.

4.8. V68T, a variant with increased heme affinity but decreased stability of the N state

The V68T mutant was selected as a variant with an unstable N apoMb state (*i.e*, increased K_{NI}^{0}) but a markedly increased affinity for heme (Hargrove and Olson, 1996). The Thr68 hydroxyl group forms an additional strong hydrogen-bond with coordinated water, which stabilizes bound hemin, reduces its rate of dissociation, and decreases K_{NH}^{0} markedly (Hargrove and Olson, 1996).

As is observed for WT metMb, the unfolding curves for V68T holoMb appear highly cooperative and indicative of a two-state mechanism, but the holoprotein unfolding transition is shifted toward a higher [GuHCl], *i.e*., ~2.0 M (Fig. 4.5). Fitting of these data to the six-state unfolding model reveals that the increase in transition midpoint is caused by an ~50 fold increase in heme affinity of the N state (K_{NH}^{0} decreases from ~0.9 to ~0.02 x10^{-13} M) even though the stability of the N apoprotein state is decreased ~5 fold (K_{NI}^{0} increases from 0.02 to 0.13, Table 4.2). The fits do suggest the presence of an IH state during unfolding, but it is present at ≤ 40% at the unfolding midpoint. We were able to deconvolute a hemichrome spectrum at this denaturant concentration, but the noise was considerable due to the smaller population of IH and larger amounts of NH and U+H states. As is the case for the other mutants, the fitted parameters suggest that the affinity of the I state for hemin is similar to that for the WT intermediate with K_{IH}^{0} ~1.7 x10^{-11} M (Table 4.4).

4.9. H97D, a variant with markedly decreased heme affinity but wild-type apoMb stability

The H97D mutant has a ~100-fold decreased affinity for hemin while retaining N and I apo-state stabilities similar to those for WT Mb. The H97D substitution disrupts favorable electrostatic interactions between the His97 side chain, the Ser92 hydroxyl group on the F-helix, and the heme-7-propionate (Liong et al., 2001). This disruption markedly increases the rate of hemin dissociation and the equilibrium dissociation constant (K_{NH}^{0}) but has little or no effect on the unfolding of apoMb (Fig. 4.2, Table 4.2 and (Hargrove and Olson, 1996)).

The unfolding curves for H97D metMb (Fig. 4.5) show inflection points, demonstrating unambiguously the presence of an intermediate, and the fitted parameters show that that the biphasic nature of the curves is caused by population of the IH state, which reaches a maximum of ~80% at ~1.5 M GuHCl. As is the case for the T67P mutant, the absorbance spectrum of H97D metMb at ~1.5 M GuHCl indicates a hemichrome species (Fig. 4.6D). The deconvoluted spectrum for the IH state of H97D Mb is shown in Fig. 4.6C, and is identical to that of the IH species for T67P and H64F Mb.

The fitted parameters for H97D holoMb show that the N state has an ~100-fold lower affinity for hemin, with K_{NH}^{0} at 84 $x10^{-13}$ M, which is similar to that estimated previously from association and dissociation, rate constants for hemin binding to this mutant (Table 4.4, Hargrove and Olson, 1996). The affinity of the I state of H97D Mb for hemin is also lower, with K_{IH}^{0} increasing from ~1.4 to ~8.8 $x10^{-11}$ M. This ~6-fold decrease in affinity suggests that the unfavorable electrostatic interactions caused by the

H97D mutation at the FG corner also weaken the affinity of the I state for hemin even though the heme pocket is unfolded.

4.10. The IH state is a hemichrome

All the Mb variants appear to populate an IH state during unfolding of the holoprotein, but the extent varies greatly from ≤ 20% for WT holoMb to ~80 and ~90% for the H97D and T67P mutants, respectively. The deconvoluted absolute spectra of the IH state for the T67P (Fig. 4.6A) and H64F (Fig. 4.6B) variants are shown and compared to the starting metMb spectra and the final U state with free hemin. The deconvoluted IH spectra of the three variants showing the highest population of IH states are overlaid in Fig. 4.6C and compared to that for H97D metMb titrated with excess imidazole (100 mM, pH 7) to create a bis-imidazole hemichrome in the NH state. All four spectra are superimposable and display the low spin peak at 535 nm, the 565 nm shoulder, and the red-shifted Soret band at 415 nm, characteristic of a bis-histidyl or a methionyl-histidyl hemichrome.

Surprisingly, the IH species appears to exhibit ~40% of the fluorescence of the completely unfolded U state even though hemin is still bound. This property is clear for T67P Mb (Fig. 4.5). At 1.4 M GuHCl, IH is the dominant species with a clear hemichrome UV-visible spectrum, roughly the same CD signal as the I state of T67P apoMb (~30% of completely folded holoMb), and roughly 40% of the fluorescence of the U state of apoMb. Because, hemin is still bound, we expected that the Trp fluorescence would remain almost completely quenched, but it is not. This result suggests that the orientations of the dipole moments for fluorescence resonance energy transfer is much less favorable than in the native state, expansion of the heme pocket size has moved the

porphyrin ring further away from the two Trp side chains positioned along the A helix, and/or the enhanced mobility of the Trp residues in this molten globule state reduces both solvent quenching and resonance energy transfer.

4.11. Reversibility of hemichrome formation in Mb

The hemichromes involved in Hb and Mb degradation have often been associated with irreversible protein degradation (Rifkind et al., 1994). However, the GuHCl-induced unfolding curves for holo-metMb are reversible and independent of how the mixtures are made. We carried out kinetic experiments in which holoMb at high [GuHCl] was rapidly diluted into buffer and *vice versa*. The same equilibrium state was obtained in each case. For example, upon dilution into buffer, the hemichrome IH state of H97D induced by 1.4 M GuHCl refolds, within the dead time of our stopped-flow apparatus, into the native metMb form with a "normal" Soret maximum at 409 nm and visible absorbance bands at 505 and 630 nm indicative of high-spin aquohemin.

We also tested whether or not the IH hemichrome in H97D could refold into a native-like pentacoordinate deoxyMb state after rapid reduction. As shown in Fig. 4.6D, when H97D metMb in 1.4 M GuHCl is reduced anaerobically with a slight excess of sodium dithionite, a "normal," pentacoordinate deoxyMb species was formed immediately and showed a broad Soret maximum at 434 nm and a single visible absorbance band centered at 558 nm. No evidence of a hemochrome (bis-imidazole complex with heme iron in the reduced state) spectrum was observed under these conditions at 1.4 M GuHCl, which Hargrove and Olson (1996) showed is too small to induce any denaturation of WT or H97D deoxyMb. Wittung-Stafshede *et al.* (1998) carried out a similar experiment with native metMb denatured with 2.5 M GuHCl. In

their experiment, reduction was induced by photo-excitation of NADH with a nanosecond Nd:YAG laser. Almost completely denatured metMb folded into pentacoordinate, native-like deoxyMb within 2-3 milliseconds ($\tau \approx 0.3$ ms) after photo-reduction. Thus, the hemichrome IH intermediate appears to be fully reversible and able to fold back rapidly into highly stable ferric or reduced NH states.

4.11. *m*-values for hemin dissociation equilibrium constants

In our analysis, we have assumed that all three hemin dissociation constants are increased in presence of [GuHCl] and that there is a linear free energy relationship between $\Delta G°$ for hemin dissociation and GuHCl concentration, as expressed by Equations 4.4-6. If K_{NH} were independent of [GuHCl], and the denaturant only affected the apoMb unfolding constants as shown in Fig. 4.2, then the computed $[GuHCl]_{1/2}$ would be ~3.5 M for WT holoMb, which is markedly greater than the observed value of ~1.8 M (Fig. 4.5). Thus, it is clear that GuHCl induces hemin dissociation, and the fitted m_{NH}, m_{IH}, m_{UH} values help define the steepness and transition midpoint of the holoMb unfolding curves. The absolute values of m_{NH} and m_{IH} are unexpectedly larger than those for the N-to-I, and I-to-U transitions of apoMb, 4.4-2.4 kcal mol^{-1} M^{-1} for hemin dissociation versus 2.3-1.4 kcal mol^{-1} M^{-1} for apoglobin unfolding.

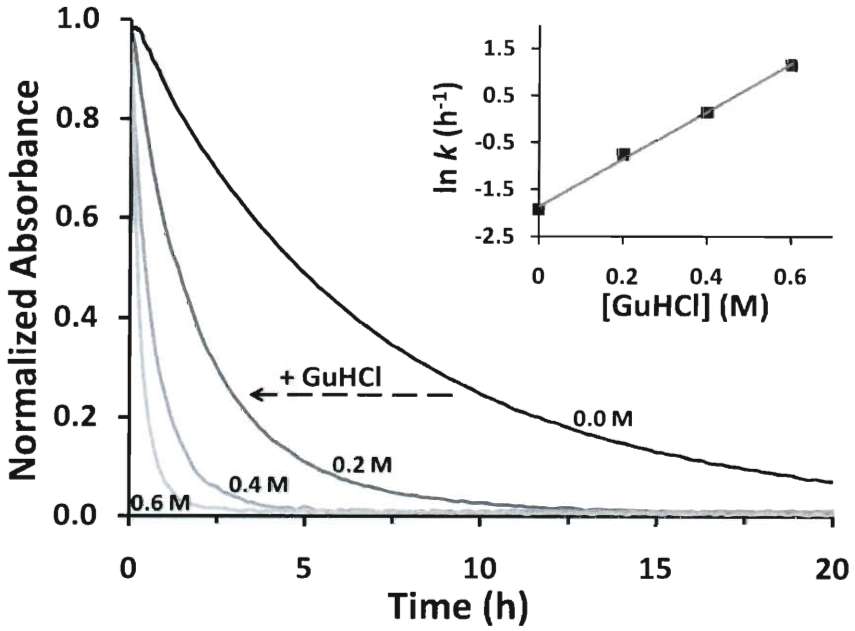


Figure 4.7. Kinetic traces for hemin loss from H97D aquomet Mb as a function of [GuHCl]. The four traces represent normalized absorbance change time courses at 409 nm as a function of GuHCl concentration, which is listed beside each curve. In these experiments, 4 μM H97D aquometMb was mixed with 40 μM H64Y/V68F apoMb, which acts as a hemin scavenging agent, in 10 mM potassium phosphate at pH 7, 20 °C (See Section 2.4). **INSET:** The observed hemin dissociation rate constants were obtained from fits to a single exponential process for each trace. A plot of the natural logarithm of these rates versus [GuHCl] is shown in the inset.

To verify the strong dependence of K_{NH} on [GuHCl], we measured the rate of hemin loss from H97D metMb as a function of GuHCl concentration in the range from 0.0 to 0.6 M, and the results are shown in Fig. 4.7. These experiments were carried out under the conditions used for equilibrium unfolding (10 mM phosphate pH 7, 20 °C) and in the absence of sucrose, which is normally added as a stabilizing agent to inhibit precipitation of the resulting apoglobin (Hargrove et al., 1994b). As result, the observed rate of hemin dissociation from H97D metMb is smaller than that reported previously for this mutant in high salt and sucrose at 37 °C (Hargrove et al., 1996b). As predicted from the results in Table 4.4, the rate of hemin dissociation depends strongly on [GuHCl], even at these low concentrations where little or no net equilibrium unfolding occurs (see 0 to

0.6 M [GuHCl] regions in Fig. 4.5). The observed rate of hemin loss from H97D metMb increases from ~0.15 h^{-1} at 0 M to ~4 h^{-1} at 0.6 M GuHCl. A plot of $\ln k_{-hemin}$ vs. [GuHCl] is linear with an apparent *m*-value for the kinetic free energy barrier to hemin dissociation of ~3.0 kcal mol^{-1} M^{-1}. This value is large and on the order of those obtained for equilibrium hemin dissociation from the NH and IH states of this mutant (*i.e.*, m_{NH} ~4 and m_{IH} ~3 kcal mol^{-1} M^{-1}, Table 4.4). The agreement between the kinetic and equilibrium m-values for hemin dissociation appears reasonable considering the difference in type of free energy measurement, *i.e.*, equilibrium $\Delta G°$ versus kinetic $\Delta G^{\ddagger}$ determinations.

4.12. Relevance of apoMb unfolding to the assembly and denaturation of the holoprotein

The unfolding of apoMb has been studied in great detail over the past 20 years. The three-state mechanism is well-established; and the structural nature of the molten globule intermediate or I state has been characterized by both mutagenesis and NMR approaches (Barrick and Baldwin, 1993b; Hughson et al., 1991; Jennings and Wright, 1993; Nishimura et al., 2006). However the relevance and applicability of these results to understand both the assembly and denaturation of holoMb had not been established until this work. The key problem is that chemical or thermal unfolding of WT holoMb is a highly concerted process, which superficially resembles a two-state process because of the high affinity of the N apoMb state for hemin (Fig. 4.5, top panels). In this case, the concentration of denaturant required to dissociate hemin from the N state is higher than that required to unfold the initial apoprotein N state and the partially unfolded molten globule I state. To resolve the role of the apoglobin intermediate in holoprotein unfolding

and assembly, we expanded on previous work (Hargrove and Olson, 1996) to include a combined analysis of the unfolding of WT and four holoMb mutants. The mutations were designed to vary the stability of the apoMb N state and its affinity for heme in order to visualize heme binding to the apoglobin folding intermediate. As shown in Fig. 4.5, this mutagenesis approach was successful. Our underlying assumption is that the same basic six-state mechanism applies to all five variants and that the mutations only affect the hemin dissociation and apoglobin unfolding parameters. The success of the fits and correlations of the parameter changes with expected structural effects imply that this assumption is a good approximation. Thus, one key conclusion is that the well-established mechanism for apoMb unfolding is directly relevant to holoMb unfolding and assembly, particularly when interpreting the effects of mutagenesis on overall stability and the appearance of intermediates.

These results are also relevant to holoprotein assembly *in vivo*. Graves *et al.* (2008) and others (Hargrove et al., 1994a; Scott et al., 2000; Smith, 2003) have suggested that expression levels of hologlobins in bacteria are proportional to the stabilities of the corresponding apoproteins because heme insertion is often limiting when high copy number plasmids are used and maximal transcription is induced. Under these conditions, the rate of holoprotein production is determined by the rate of heme synthesis or transport times the fraction of apoglobin that is folded and competent to bind heme (Graves et al., 2008). One key question in this interpretation is whether heme can bind to the molten globule apoglobin intermediate and facilitate formation of the holoprotein.

4.13. The IH hemichrome state and holoMb assembly

Hemichromes have long been associated with irreversible degradation of Hb and formation of Heinz bodies in red cells (Rifkind et al., 1994), and presumably similar degradation processes occur for Mb *in vitro* or in myocytes. Our results in Figs. 4.5 and 4.6 show that sw Mb unfolds via an IH state, which displays the spectrum of a hemichrome with a bis-histidyl linkage. The observation of a hemichrome spectrum for the H64F variant demonstrates that the distal histidine is not needed to produce the low spin IH complex. Thus, the hemichrome spectrum is probably due to a mixture of bis-imidazole or methionyl/histidyl complexes involving any two of the histidines located in the region of the heme pocket, His24(B5), His36(C1), His48(CD6), His64(E7), His81(EF corner), His82(EF loop), His93(F8), and His97(FG loop) or one of these histidines and Met55(D5) (Fig. 4.1). The proposed flexibility of the molten I and IH states appears to allow different combinations, although the simplest structural interpretation for most of the variants would be axial ligation by the distal and proximal histidines (His64 and His93) when they are both present.

Reversible hemichrome formation is a general phenomenon and is observed during either GuHCl or urea-induced denaturation of both sw and horse heart Mb (not shown). Hemichrome formation is fully reversible, and the IH state reverts to the native metMb form with distal water coordinated to the iron atom upon rapid dilution of GuHCl. Similarly, the IH hemichrome does not form a hemochrome when reacted with dithionite. Instead, the IH state rapidly converts to a native pentacoordinate deoxyMb form upon reduction (Fig. 4.6D), and previous laser induced reduction experiments suggest that this process occurs in milliseconds (Wittung-Stafshede et al., 1998). Our equilibrium results

suggest that hemin can bind with reasonably high affinity to the I state, and promote formation of the native holoprotein. This pathway for assembly is even more probable under micro-aerobic conditions similar to those found in *E. coli* growing in logarithmic phase and, presumably, in respiring myocytes. Under these conditions, free heme still rapidly autooxidizes, but reduction of the ferric IH state leads to rapid formation of very stable deoxy- or oxy-holoMb states.

4.14. Hemin binding to the N, I, and U apoMb states

It was difficult to determine the N and I state hemin dissociation equilibrium constants accurately without fixing the value of the equilibrium constant for hemin dissociation from the unfolded state. K_{UH}^{0} is poorly defined in our experiments at 10 μM initial holoMb concentration because no significant amount of UH occurs at high [GuHCl] (Fig. 4.5). Separate gel-filtration chromatography experiments suggest complete hemin dissociation upon unfolding of ~10 μM metMb at [GuHCl] ≥ 3 M (not shown). Spectra of the Mb variants equilibrated in 3 M GuHCl resemble that of free hemin in 3 M GuHCl. We fixed K_{UH}^{0} at 1.0 x10^{-6} M in our model based on the affinity constants measured for non-specific hemin binding to unfolded apo-cytochrome b_{562} (Robinson et al., 1997) and native bovine serum albumin (Hargrove et al., 1996a; Marden et al., 1989) and on the estimation that non-specific hemin binding to unfolded apoMb is ~10^{6} to 10^{7}-fold weaker than binding to the folded N apo-state (Hargrove and Olson, 1996). Fixing K_{UH}^{0} to 10^{-6} M, and K_{NI}^{0} and K_{IU}^{0} to the values determined in separate experiments with the apoMb variants, allowed determination of the equilibrium dissociation constants for hemin binding to the N and I apoglobin states.

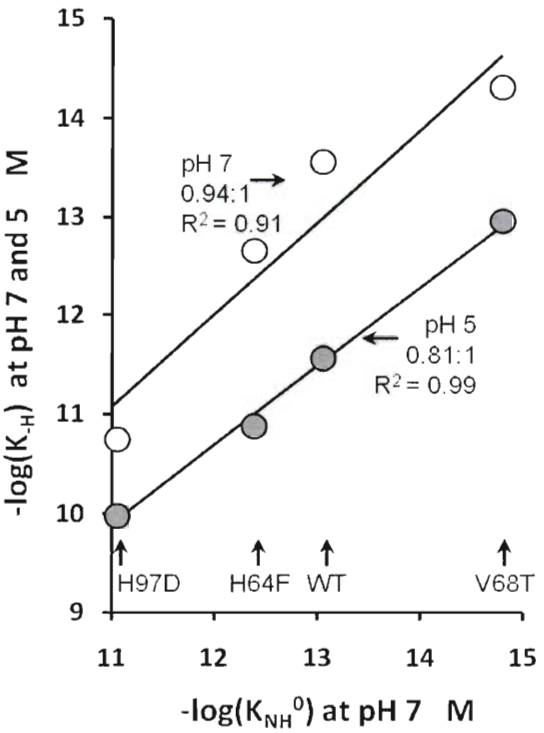


Figure 4.8. Correlation of hemin affinities. The correlations are between the hemin dissociation equilibrium constant, K_{NH}^0, obtained from the analysis of the GuHCl unfolding curves shown in Fig. 4.5 with hemin dissociation equilibrium constants, K_{-H}, calculated from the ratio of the association and dissociation rates constants measured independently in kinetic experiments (Hargrove et al., 1996a; Hargrove and Olson, 1996; Hargrove et al., 1996b). The rate parameters are listed in Table 4.4. The open circles represent K_{-H} values at pH 7-8 and the closed, gray circles, K_{-H} values at pH 5. The solid lines represent linear fits for $-\log(K_{NH}^0)$ versus $-\log(K_{-H})$ with the slopes and R^2 values listed beside each set of data. The K_{NH}^0 values were measured in 0.01 M potassium phosphate, pH 7, 20 °C. Most of the K_{-H} values were obtained from measurements in 0.15 M sodium phosphate, 0.45 M sucrose, pH 7.0 or 0.15 M acetate, 0.45 M sucrose, pH 5 both at 37 °C. The correlations are very strong, despite the differences in conditions.

The results in Tables 4.1 and 4.2, and the fits in Fig. 4.5 show that distal pocket mutations affect primarily stability of the N state and its affinity for hemin. The variations in K_{NI}^0 and K_{NH}^0 (~1,000-fold) are markedly greater than those for K_{IU}^0 and K_{IH}^0 (~6-fold). These results strongly verify Baldwin's, Barrick's, Wright's, and coworkers' interpretation (Barrick and Baldwin, 1993b; Hughson et al., 1991; Jennings and Wright, 1993; Nishimura et al., 2006) that the N-to-I, and NH-to-IH transitions involve unfolding of the heme pocket to generate a molten globule intermediate with intact A, G, and H helices.

Another key result is that the variation in K_{NH}^0 values determined by analysis of holoMb unfolding curves correspond well with equilibrium hemin dissociation constants, K_{-H}, calculated from the ratio of dissociation and association rate constants for hemin

binding to the same set of Mb mutants (Table 4.2, last two columns, (Hargrove et al., 1996a; Hargrove and Olson, 1996; Hargrove et al., 1996b)). As shown in Fig. 4.8, there is a strong, linear correlation between -log(K_{NH}^{0}) obtained independently from unfolding curves and the values of -log(K_{-H}) determined from kinetic experiments at both neutral and low pH. The neutral pH values of K_{-H} are poorly defined because the rates of hemin dissociation from WT metMb and the other higher affinity mutants are very small and difficult to measure. At pH 5, hemin dissociation is more readily measured because of increased protonation of the proximal histidine (His93) which promotes breakage of the Fe-imidazole bond (Hargrove et al., 1996b). Despite the differences in solvent conditions and temperature, both the absolute values of K_{NH}^{0} and the effects of mutagenesis correlate remarkably well with the K_{-H} values for WT, H64F, V68T, and H97D, determined kinetically at both pH values (Table 4.4 and Fig. 4.8). Thus, GuHCl-induced unfolding curves of holo-heme proteins represent an alternative assay to measure accurately the affinities of apoproteins for hemin if the mechanism for unfolding of the apoprotein is known.

Chapter 5

Influence of hemin dimerization and protein concentration on the unfolding of holomyoglobin

5.1. Introduction

Resolving how native holoMb folds and assembles into its native structure is more complex than anticipated and requires taking into account: the folding mechanism of the apo-form, heme binding to the different apoMb states, the redox state of heme and the associated self-associating properties of the free heme cofactor in aqueous solutions. Most previous studies have presumed that holoMb unfolds by a one step-transition/two-state mechanism, despite the presence of a least one intermediate in the folding pathway of the apo-form. In Chapter 4, we examined the role of heme in the unfolding of holoMb using a set of well chosen Mb variants with properties affecting either the stability of the apo-form or hemin binding (Culbertson and Olson, 2010). Our results presented in Chapter 4 support previous studies that estimated the affinity of the fully folded apo-state for hemin to be ~10^{-14} M (Gibson and Antonini, 1960; Hargrove et al., 1996a; Hargrove et al., 1996b). More importantly, we were able to estimate the affinity of the apo-intermediate (I) state for hemin to be ~10^{-11} M, as the unfolding of the met-holo form does not proceed via a simple two-state mechanism but rather via a reversible hemin-bound intermediate which has the spectral characteristics of a hemichrome species (Soret band at ~ 416 nm, and low spin band at 535 nm with a shoulder at 565 nm).

Hargrove and Olson (1996) conducted equilibrium-unfolding experiments on WT holoMb and found very little dependence on protein concentration contrary to what would be expected if heme were released upon unfolding of the polypeptide chain (Model 1, Chapter 3, Fig. 3.1A). They attributed the lack of protein concentration dependence to non-specific binding of heme to the unfolded polypeptide. The lack of protein concentration dependence on apparent stability was also found with ferric cytochrome b_{562} which showed no significant increase in thermal stability upon increasing protein concentration (Robinson et al., 1998).

It has been known for a long time that hemin (ferric heme) has a propensity to self-associate, and recent publications have demonstrated that hemin can form different π-π stacking or μ-oxo dimers depending on the conditions (Asher et al., 2009; de Villiers et al., 2007). Thus, the dimerization properties of hemin could also modulate the equilibrium unfolding of holo-globins, as shown in Chapter 3 (Model 3, Fig. 3.1C).

The effects of protein concentration and the possible influence of hemin dimerization are examined in this Chapter for the apo and holo forms of the WT and H97D Mb. The H97D variant has a decreased affinity for hemin in the native folded state due to the removal of a hydrogen bond network between the imidazole side chain of His97, Ser92, and the heme 7–propionate (Hargrove and Olson, 1996; Liong et al., 2001), and strongly populates the IH state during equilibrium unfolding (Fig. 4.5, Culbertson and Olson, 2010). Our goal was to determine if a comprehensive analysis of these two Mbs could help resolve whether hemin dimerization or hemin binding to the unfolded state was the major cause of the lack of protein concentration dependence on holoMb unfolding. The expressions for analyzing the holoprotein unfolding curves should apply

generally to all hemoproteins and most proteins with relatively insoluble cofactors. The details of the analyses build upon conclusions from the simple models in Chapter 3 and the complex six-state mechanism described in Chapter 4.

5.2. Equilibrium-unfolding of WT and H97D apoMb variants

As in Chapter 4, the analysis of apoMb unfolding is based on a three-state model of the apoMb unfolding (Barrick and Baldwin, 1993a; Barrick et al., 1994; Culbertson and Olson, 2010; Griko et al., 1988; Nishii et al., 1995; Scott et al., 2000). In this work, the buffer concentration was changed from 10 to 100 mM phosphate to mimic more closely the cellular conditions (ionic strength and osmolarity). The stability parameters K_{NI}^{0} and K_{IU}^{0} for the N-to-I and I-to-U transitions of the apoMb variants were needed for analysis of the holoprotein unfolding. Thus, both WT and H97D apoMb variants were titrated with GuHCl, and the unfolding was followed by far-UV CD and fluorescence emission. As in Chapter 4, the CD data at 222 nm alone do not allow for an accurate measure of the amount of intermediate present. The latter species is revealed by the fluorescence emission data at 343 nm as the hyperfluorescence peak in the bell-shaped curve. The wavelength at which maximal emission intensity occurs is red shifted by 2 nm as a result of the increase in phosphate concentration. The unfolding data from both detection data were analyzed using a global fit, again assuming a simple three-state scheme.

The equilibrium unfolding curves were successfully fit to the three-state equilibrium unfolding mechanism (Fig. 5.2) and the K_{NI}^{0} and K_{IU}^{0} stability constants for the N-to-I and I-to-U transitions are shown in Table 5.1. As expected, the H97D mutation does not significantly affect the stability of the N and I states of apoMb, as His97 is a

surface residue, already solvent exposed in the folded apo form. The increase in phosphate concentration from 10 to 100 mM did however increase the overall stability of the two apoMb variants by a factor of ~2 to 3 (as stability constants) which corresponds to a free energy increase of 0.4 to 0.6 kcal mol^{-1}. Phosphate and potassium ions are part of the Hofmeister series, which are known to decrease the solubility of hydrophobic residues and therefore strengthen hydrophobic interactions (Baldwin, 1996), which in apoMb help the hydrophobic core to resist denaturation. This effect was also noticed in a previous study on apoMb variants (Smith, 2003), and appears to be relevant to all apoglobins. The 10-fold increase in buffer concentration also affected the fluorescence intensities of the intermediates for both (Figs. 4.2 and 5.1).

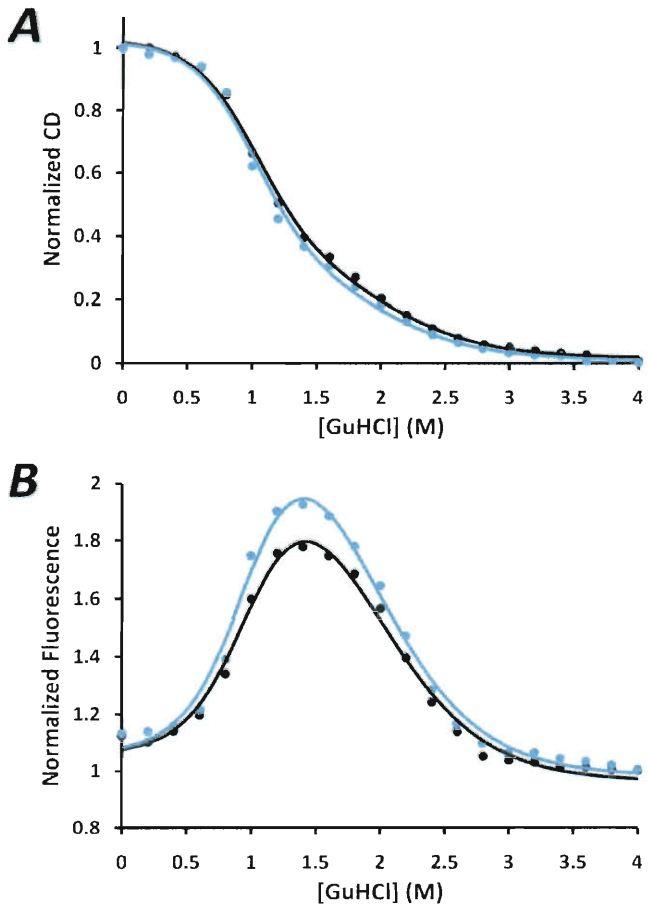


Figure 5.1. Equilibrium-unfolding of WT and H97D sw apoMb. The unfolding curves obtained by normalizing the changes in far-UV CD at 222 nm to 1.0 for native WT apoMb and assuming that the U states of all five variants have the same CD_{222nm} (panel A) and by normalizing fluorescence emission intensities at 343 nm to 1.0 for the final, completely unfolded state of each mutant (panel B). Global fits to a three-state mechanism are shown for WT (black) and H97D (cyan). Conditions: 10 μM protein, 100 mM potassium phosphate buffer pH 7, 20 °C.

Table 5.1. Stability parameters for equilibrium unfolding of apoMb variants[1]

ApoMb variant	K_{NI}^0	K_{IU}^0	$1/K_{NI}^0 K_{IU}^0$
WT	0.016	0.011	5400
H97D	0.017	0.012	4800

[1] The GuHCl-induced unfolding measurements were carried out in 100 mM potassium phosphate pH 7, 20 °C. The K_{NI}^0 and K_{IU}^0 were obtained by fitting simultaneously both CD and Fluorescence signals for the unfolding of the apoMb variants to the three-state unfolding mechanism using Equation 4.3 (Appendix A.4). The values of m_{NI} and m_{IU} were set to 2.35 and 1.36 kcal mol^{-1} M^{-1}, respectively.

5.3. A six-state mechanism with hemin dimerization for the unfolding of holoMb

The equilibrium-unfolding mechanism described in Chapter 4 was used to investigate the unfolding of WT and H97D Mb variants in the aquomet form with the addition of hemin self-association and variation of protein concentration.

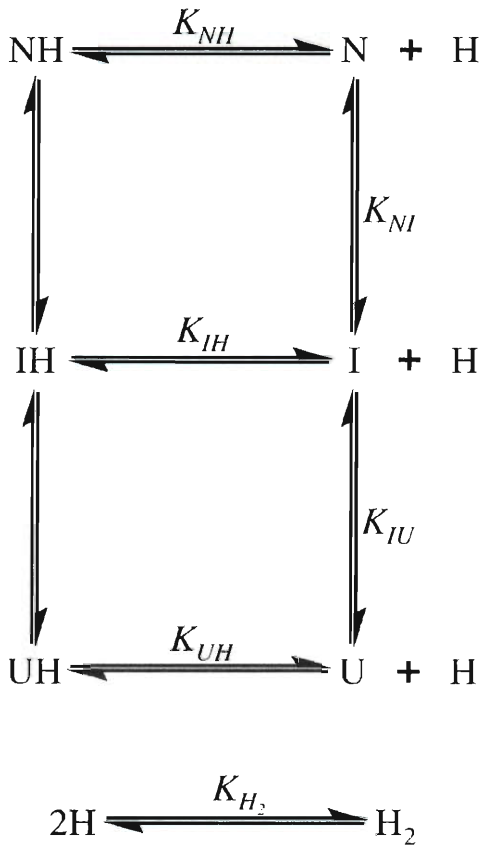


Figure 5.2. Six-state model for the unfolding of holoMb based on the three-state unfolding mechanism of apoMb, the assumption that heme can bind to all three apoMb states, and that heme self-dissociates.

The model shown in Fig. 5.2 was applied to the unfolding curves for each holoMb variant. In this equilibrium model, the K_{NI}^{0} and K_{IU}^{0} stability constants and m_{NI} and m_{IU} values for the N-to-I, and I-to-U transitions of apoMb were determined independently in separate experiments with the apo-forms. The heme dissociation constants from all possible heme-bound states, NH (metMb), IH, and UH were allowed to vary in order to define the equilibrium unfolding curves for holoMb. The hemin dimerization constant is defined as:

$$K_{H_2} = \frac{[H_2]}{[H]^2} = K_{H_2}^{0} \exp\left(\frac{m_{H_2}[GuHCl]}{RT}\right) \tag{5.1}$$

The population fractions of each species were computed by first obtaining the free concentration of heme [H], which can be computed by finding the appropriate root of the following cubic function:

$$a[H]^3 + b[H]^2 + c[H] + d \tag{5.2}$$

where:

$$a = 2K_{H_2}\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)$$

$$b = 2K_{H_2}(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}) + \left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)$$

$$c = K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}$$

$$d = -P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right) \tag{5.3}$$

The derivation of Equation 5.3 is given in the Appendix A.6. The fraction of each species was calculated using the free heme concentration [H], the two apoMb stability constants, and the three hemin equilibrium dissociation constants as shown beneath for the NH state:

$$Y_{NH} = \frac{1}{1+\frac{K_{NH}}{[H]}+\frac{K_{NH}K_{NI}}{[H]}+\frac{K_{NH}K_{NI}K_{IU}}{[H]}+\frac{K_{NH}K_{NI}}{K_{IH}}+\frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}} \quad (5.4)$$

Each of the stability and affinity constants in Equation 5.4, were computed at different [GuHCl] using the *m*-values: m_{NI}, m_{IU}, m_{NH}, m_{IH}, and m_{UH}, as defined in Equations 4.1-2 and 4.4-6 to determine the fractions of each species as a function of [GuHCl].

The Soret absorbance and far-UV CD signals were computed from:

$$S = S_N Y_N + S_{NH} Y_{NH} + S_I Y_I + S_{IH} Y_{IH} + S_U Y_U + S_{UH} Y_{UH} + S_H Y_H + S_{H2} Y_{H2} \quad (5.5)$$

The two detection methods, CD, and Soret absorbance, used for analysis of the equilibrium-unfolding of ferric holoMb variants, were analyzed at single wavelengths, 222, and 409 nm, respectively. Unfolding curves were determined at three protein concentrations, 2, 10, and 100 μM. The use of fluorescence detection was not possible due to the high range of protein concentrations and subsequent absorbance at 280 nm. The unfolding data sets for the two detection methods at the three protein concentrations were simultaneously fit using an algorithm coded in C for use with Gnuplot (http://www.gnuplot.info/) and based on the unfolding mechanism in Fig. 5.2. The full source code is provided in the Appendix A.10.

5.4. Equilibrium-unfolding of WT and H97D holoMb variants

Even though WT aquomet holoMb appears to unfold in an apparent two-state equilibrium mechanism at neutral pH, we have shown that a small amount of a hemichrome IH state is populated near the GuHCl midpoint at 10 μM (Chapter 4, Fig. 4.5, and Culbertson and Olson, 2010). The H97D mutation was designed to demonstrate unequivocally the presence of a hemin-bound intermediate even at low protein concentration (P_0= 10 μM). The difference in the unfolding curves of WT and H97D ferric Mb (Fig. 4.5) do not arise from the unfolding properties of the apo forms but from the large decrease in hemin affinity of the H97D mutant when compared to WT Mb (See Chapter 4).

The equilibrium-unfolding data for both CD and Soret absorbance detections at the three protein concentrations for WT holoMb are shown in Fig. 5.3. The data were fit using a comprehensive algorithm to interpret simultaneously the six unfolding curves. We estimated the K_{NH}^{0} and K_{IH}^{0} values to be 7.4 x10^{-14} and 1.2 x10^{-11} M, respectively. The K_{NH}^{0} and K_{IH}^{0} values for the H97D variant were estimated to be 1.5 x10^{-12} and 1.5 x10^{-11} M (Fig. 5.4). As expected, the fitted dissociation constants for hemin dissociation from the NH and IH states were little affected by free hemin dimerization when compared to the values reported in Chapter 4 where free hemin was assumed to remain monomeric.

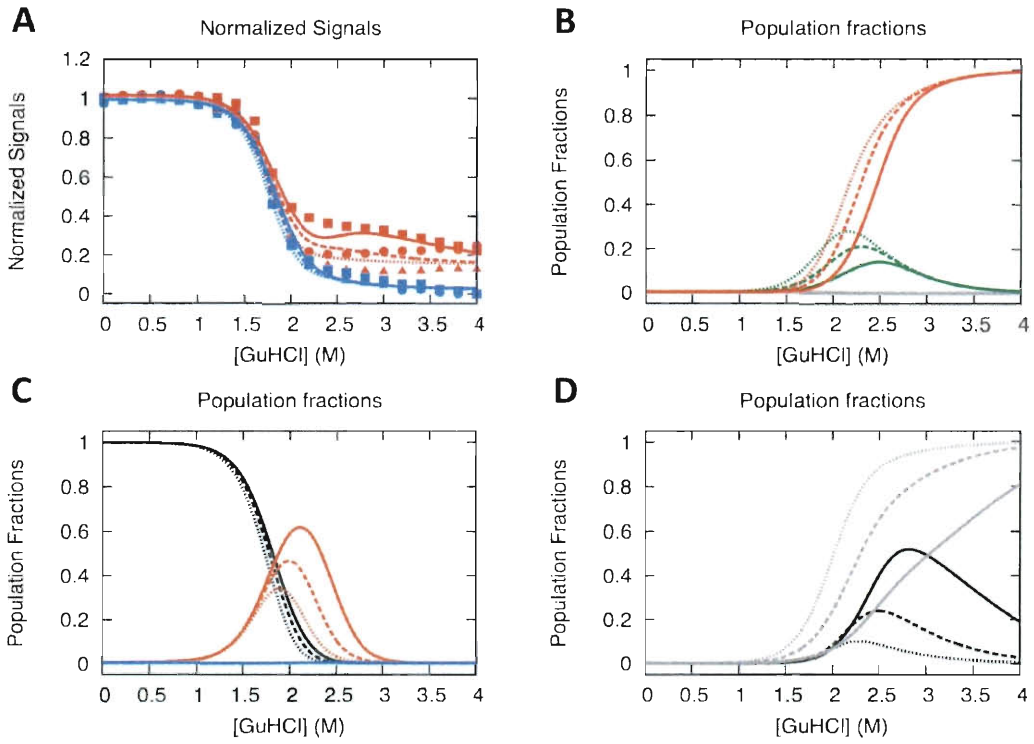


Figure 5.3. Equilibrium unfolding of WT sw holoMb at different globin concentrations. **A.** Unfolding data for CD (blue) and Soret absorbance (red) data at 2 (round), 10 (triangle), and 100 (squares) μM globin concentrations. The best fits to the 2 (dotted lines), 10 (dashed lines) and 100 (solid line) μM protein data are shown. **B.** Population fractions of the N (gray), I (green) and U (red) apo states. **C.** Population fractions of the NH (black), IH (orange) and UH (blue) holo states. **D.** Population fractions of H (gray) and H_2 (black). The CD data is depicted in blue, the Soret absorbance in red. Conditions: 100 mM potassium phosphate pH 7, 20 °C.

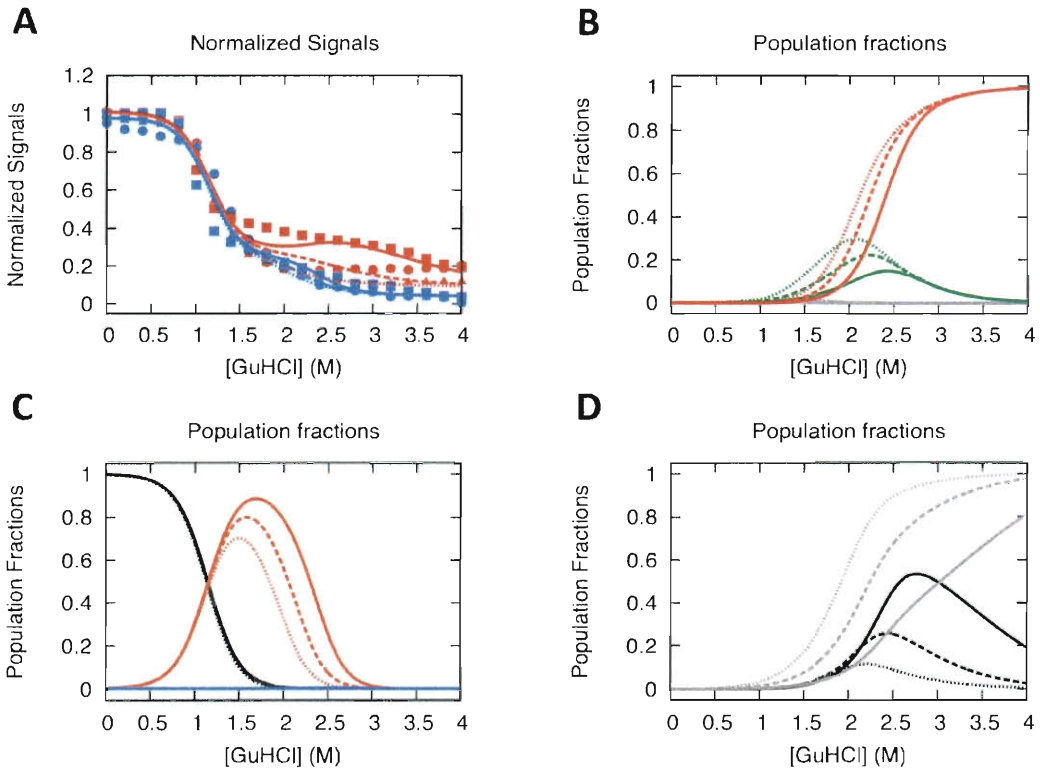


Figure 5.4. Equilibrium unfolding of H97D sw holoMb at different globin concentrations. **A.** Unfolding data by CD (blue) and Soret absorbance (red) at 2 (round), 10 (triangle), and 100 (squares) μM globin concentrations. The best fits to the 2 (dotted lines), 10 (dashed lines) and 100 (solid line) μM are shown. **B.** Population fractions of the N (gray), I (green) and U (red) apo states. **C.** Population fractions of the NH (black), IH (orange) and UH (blue) holo states. **D.** Population fractions of H (gray) and H_2 (black). The CD data is depicted in blue, the Soret absorbance in red. Conditions: 100 mM potassium phosphate pH 7, 20 °C.

The new fits reveal that increasing protein concentration favors the IH state while decreasing population of the apo-I state. Even though the fits are not perfect, this analysis does show that the lack of protein concentration dependence of the unfolding curves is mainly due to hemin binding to the I state and that the "extra" Soret absorbance phase observed at high [GuHCl] is probably due to hemin dimerization after protein unfolding is complete. The deviations of the fits to the Soret data are probably a result of uncertainties regarding the exact aggregation state of hemin in GuHCl solutions.

5.5. Effect of protein concentration on formation of the IH hemichrome and I states

In principle, hemin dissociation should be favored a low concentrations, and increasing P_0 would favor the hemin bound states, NH, IH, and UH. At high [GuHCl], the U state does not bind hemin with high enough affinity to retain hemin at spectroscopically detectable levels even at high protein concentration. However, the I state can specifically bind heme and the IH hemichrome state is readily detectable and markedly affects the measured unfolding curves. Increasing protein concentration favors the IH state and enhances resistance to hemin loss and complete unfolding. The net result is that the hemichrome state is much more readily populated and detected at high globin concentrations.

The unfolding curves for both WT and H97D holoMb variants are shown in Fig. 5.3, and as predicted by the scheme in Fig. 5.2, the presence of the IH state is significantly more favored at 100 μM than at 10 or 2 μM total protein. In addition, these analyses show that the population of the IH state is greater at the higher 100 mM

potassium phosphate concentration used in these experiments than the 10 mM concentration used in Chapter 4. This effect is particularly evident for WT holoMb.

The effect of the denaturants GuHCl or urea on the self-association properties of hemin have not been reported previously. Our preliminary results using the absorbance characteristics of hemin suggest that hemin dimerization is little affected in solutions containing chemical denaturants (GuHCl and urea). Equation 5.1 assumes that GuHCl exponentially influences the hemin self-association constant. In Chapter 3, we investigated the possible cause for the lack of protein concentration dependence on the apparent stability of monomeric globins. The addition of a hemin dimerization step alone could explain part of the lack of protein concentration dependence of the unfolding curves of globins (Chapter 3, Fig. 3.1C). GuHCl-induced hemin self-association (*i.e.*, a positive m_{H_2} value) would facilitate dissociation of hemin at high protein concentration and compensate the expected stabilization of the NH and IH states at high holoMb levels. The net result of a positive m_{H_2} value would be less protein concentration dependence on the apparent unfolding curves. In contrast, a negative m_{H_2} value would increase the protein concentration dependence. The fits in Figs. 5.4-5 suggest a negative m_{H_2} value (~ -1.2 kcal mol^{-1} M^{-1}) for hemin dimerization so that at high [GuHCl], free hemin remains monomeric. However, the effect of [GuHCl] on hemin dimerization is small and not a major determinant of the holoMb unfolding curves except for the absorbance changes associated with the hemin dimerization at high holoMb concentrations.

5.6. Lack of apparent protein concentration dependence on equilibrium unfolding of monomeric globins

Varying protein concentration is a useful tool in analysis of the folding of multimeric or cofactor-proteins. Mathematically, increasing the concentration of a multimeric protein stabilizes it against dissociation and unfolding as illustrated by Equation 1.1. Cofactor-proteins also gain stability against dissociation of the cofactor and unfolding at elevated protein concentration.

In Mb, the lack of protein concentration dependence on apparent stability was first observed by Hargrove and Olson (1996) and was attributed to hemin remaining bound non-specifically to the unfolded polypeptide. To test this assumption, I used a gel filtration assay to determine whether hemin would separate from holoMb, unfolded in 3 M GuHCl (data not shown). Despite hemin eluting separately from unfolded apoMb, the results were inconclusive as hemin can also interact readily with sugar polymers that make up the gel filtration media. However, the assay did prove that the hemin affinity of the U state is quite weak. The unfolding simulations in Fig. 3.1 showed that hemin dimerization upon dissociation from the unfolded globin could also be a governing factor for the lack of protein concentration dependence on apparent stability. It is unlikely that hemin remains attached to unfolded Mb *in vitro* at high denaturant concentrations, as GuHCl markedly weakens the affinity of the apoMb states for hemin. It was shown that a fly-casting mechanism is probably not governing the assembly of Mb *in vivo* because folding is not accelerated in the presence of dicyano-hemin *in vitro* (Crespin et al., 2005).

Our analyses show that it is important to understand how denaturants influence the affinity of a protein for a cofactor to avoid bias in analyses of unfolding data. The

dependence of hemin affinity on [GuHCl] is shown in Chapter 4, Fig. 4.7, and the fits to unfolding data for holoMb variants in Fig. 4.5 show that the unfolded states do not possess high enough affinity to retain hemin at [GuHCl] favoring complete unfolding. Hence, it is unlikely that hemin binds to the U state prior to folding, but the results in Figs. 4.5 and 5.3-4 demonstrate that the apoMb intermediate possesses a hemin affinity high enough to occur during refolding, and it seems very likely that the IH state is probably on pathway for assembly *in vivo*. Thus the lack of protein concentration dependence on the observed unfolding curves for holoMb is primarily due to hemin binding to the I state (Figs. 5.4-5).

5.7. Effect of potassium phosphate concentration on unfolding of hologlobins.

Our previous analysis was done at low buffer concentration. In this work, the phosphate concentration was increased to 100 mM to mimic the ionic strength and osmolarity in cells. To investigate the effect of buffer concentration more directly, the dependence of the rate of hemin loss from aquomet H97D sw Mb on potassium phosphate concentration was measured (Fig. 5.5). Interestingly, increasing the buffer concentration from 10 to 300 mM accelerated the rate of hemin loss more than 2-fold.

As described earlier in this chapter, the increase in phosphate concentration stabilizes the apoMb variants, decreasing $K_{NI}^{0}K_{IU}^{0}$ by a factor of 2 to 3. If we consider the stability of the NH state to be roughly represented by $1/(K_{NH}^{0}K_{NI}^{0}K_{IU}^{0})$, it appears that the stabilizing effects on the apoprotein with increasing ionic strength are almost exactly compensated by a weakening effect on hemin affinity. More experiments are needed to

determine if this is specific to phosphate or simply an ionic strength effect.

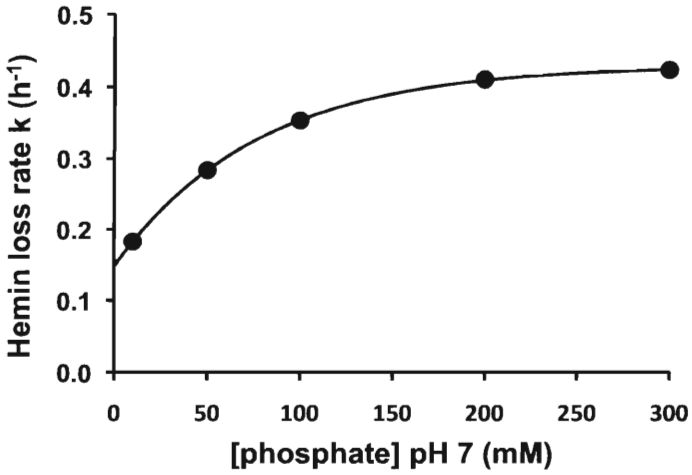


Figure 5.5. Dependence of [phosphate] on rates of hemin loss from H97D sw Mb. Hemin loss was induced upon addition of the apo reagent (H64Y/V68F Mb) and the concentration of phosphate was varied from 10 to 300 mM at pH 7, 20 °C.

5.8. Conclusions

The addition of a hemin dimerization step in the unfolding model for holoMb renders the analysis of unfolding data deceivingly much more complex. Investigating unfolding at three protein concentrations allows for a more accurate determination of hemin affinity constants. Due to the uncertainties in spectral assignments for the free hemin aggregates, it was difficult to accurately fit the last phases of the unfolding curves. However, the major phases of these curves were readily analyzed to accurately estimate the affinity constants. This detailed analysis demonstartes that hemin binding to the I state is the main cause for to the lack of protein concentration dependence on apparent stability of holoMb.

Chapter 6

Cerebratulus lacteus Hb, a simpler system

6.1. Introduction

As a control for the mechanistic approach developed for sw Mb, we looked for a simpler model globin system. The neuronal miniglobin from *Cerebratulus lacteus* (termed CerHb) was chosen for several reasons. First, this small globin lacks an A-helix, which is one of the main structural components of the molten globule intermediate found in sw apoMb. Second, the stability and folding characteristics of CerHb had not been measured previously. Third, our laboratory already had well-established protocols for expressing and purifying large amounts of CerHb. Fourth, this miniglobin should be a simpler system, in which the unfolding of the apo- and holoprotein could be interpreted using one of the simple unfolding mechanisms presented in Chapter 3. The key was to test the generality of the analysis developed for holoMb unfolding and perhaps establish a simpler model system for globin/heme assembly.

6.2. Stability of the WT apoCerHb

The extraction of the heme cofactor from ferric CerHb and proper refolding of the apoprotein was possible using the methyl ethyl ketone method (See Chapter 2). Then, WT apoCerHb was titrated with GuHCl at 20 °C as described previously for apoMb variants in Chapters 4 and 5 (Fig. 6.1). WT apoCerHb in the folded state does possess a typical α-helical far-UV CD spectrum, but has about half of the molar ellipticity of

apoMb. Fluorescence measurements of the apo form rely on emission form the Trp3 (at the beginning of the B helix) and four Tyr residues that are mostly buried in the folded state. Upon addition of GuHCl, fluorescence emission decreases due to unfolding and solvent exposure of the fluorophores, and the emission peak shifts from 333 to 354 nm.

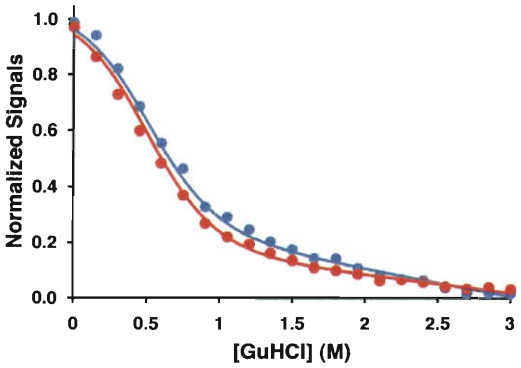


Figure 6.1. GuHCl-induced equilibrium unfolding of WT apo-CerHb, monitored by Far-UV CD_{222} (blue), and Trp fluorescence at 333 nm (red). Lines represent the best fits to the a two-state unfolding mechanism with a post-transitional baseline. 100 mM phosphate pH 7, 10 μM, 20 °C.

WT apoCerHb does unfold by a two-state mechanism as suggested by an overlap of the fluorescence and CD curves, and the lack of an inflection point. Fitting to a two-state equilibrium mechanism (N to U) with a post-transitional baseline showed that apoCerHb is quite unstable at 20 °C, with a K_{NU}^{0} of 0.127, and a *m*-value for the unfolding, m_{NU}, of 2.46 kcal mol^{-1} M^{-1}. This m_{NU} value is smaller than the combined *m*-values for the unfolding transitions of apoMb, which is in agreement with its decreased size (Auton et al., 2007).

6.3. Four-state analysis of the equilibrium unfolding of holoCerHb

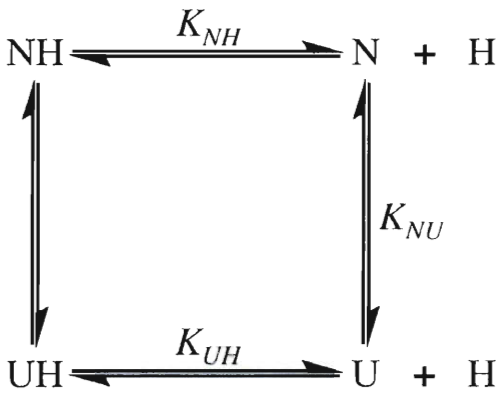


Figure 6.2. Four-state mechanism for the unfolding of holoCerHb. K_{NU} is the stability constant between the N and U apo states, and K_{NH} and K_{UH} represent the equilibrium constants for hemin dissociation from the native NH and UH states, respectively.

The four-state mechanism shown in Fig. 6.2 was used to investigate unfolding of WT holoCerHb. The equations for this mechanism are described in Chapter 3 and the appendix (Fig. 3.1B, A.2). Unfolding of WT aquometCerHb was induced by addition of GuHCl and monitored by far-UV CD, and Soret absorbance. The overlays of the unfolding curves are shown in Fig. 6.3. Fits to the unfolding data estimate the equilibrium dissociation constant for hemin dissociation from the apo-N state, K_{NH}^{0}, equal to be ~ $1.2\text{x}10^{-10}$ M. In addition, the m_{NH} value was estimated to be ~6.2 kcal mol^{-1} M^{-1}, higher than the estimated m_{NH} values for the sw Mb variants in Chapter 4 (Table 4.2).

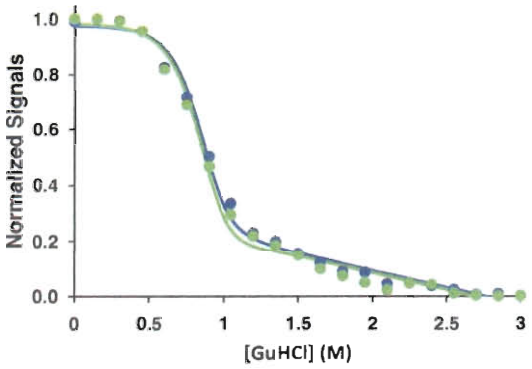


Figure 6.3. GuHCl-induced equilibrium-unfolding of WT holoCerHb, monitored by far-UV CD_{222} (blue) and Soret absorbance at 406 nm (green). Lines represent the best fits to the four-state unfolding model. K_{NH}^{0}= $1.2\text{x}10^{-10}$ M and m_{NH}= 6.2 kcal mol^{-1} M^{-1}. Conditions: 10 μM globin and 100 mM phosphate pH 7 at 20 °C.

6.4. Dependence of [GuHCl] on the rates of hemin loss from WT aquomet CerHb

The large dependence of GuHCl on the rates of hemin loss from aquomet Mb is described in detail in Chapter 4. We tested whether these high m_{NH} values are inherent to the release of hemin from all monomeric globin by examining the dependence of rate of hemin loss from metCerHb on [GuHCl]. The absorbance spectrum for aquomet CerHb is similar to that of WT aquomet Mb, which again allowed the use of H64Y/V68F apoMb (See Chapter 2.2) to displace hemin from ferric CerHb and measure the rate of hemin loss from the invertebrate Hb. As expected, WT CerHb loses hemin at a much greater rate (~0.2 h^{-1}), 20 °C, than WT Mb (~0.01 h^{-1}) at 37 °C (Hargrove et al., 1996b), and hemin loss is greatly accelerated at higher temperatures. The absolute value of k_{-h} is similar to that from H97D Mb variant (~ 0.15 h^{-1}) at 20 °C. GuHCl dramatically increases the rate of hemin loss from CerHb as in Mb, and the apparent m_{NH} value is large and comparable to that of H97D sw met-Mb. Thus, the high values of m_{NH} appear to be inherent to hemin dissociation from all globins.

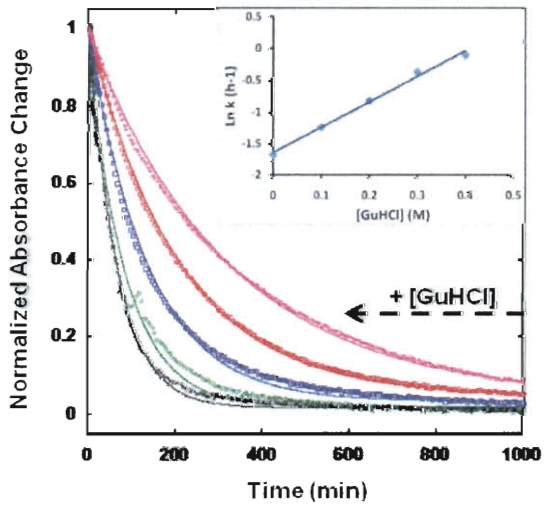


Figure 6.4. GuHCl-dependence of hemin loss rates from 4 μM WT aquomet CerHb using 40 μM H64Y/V68F apoMb. Main panel depicts the normalized absorbance decrease at 406 nm in 0, 0.1, 0.2, 0.3, and 0.4 M GuHCl. The inset depicts the natural logarithms of the rates as a function of GuHCl. 100 mM phosphate pH 7, 20 °C.

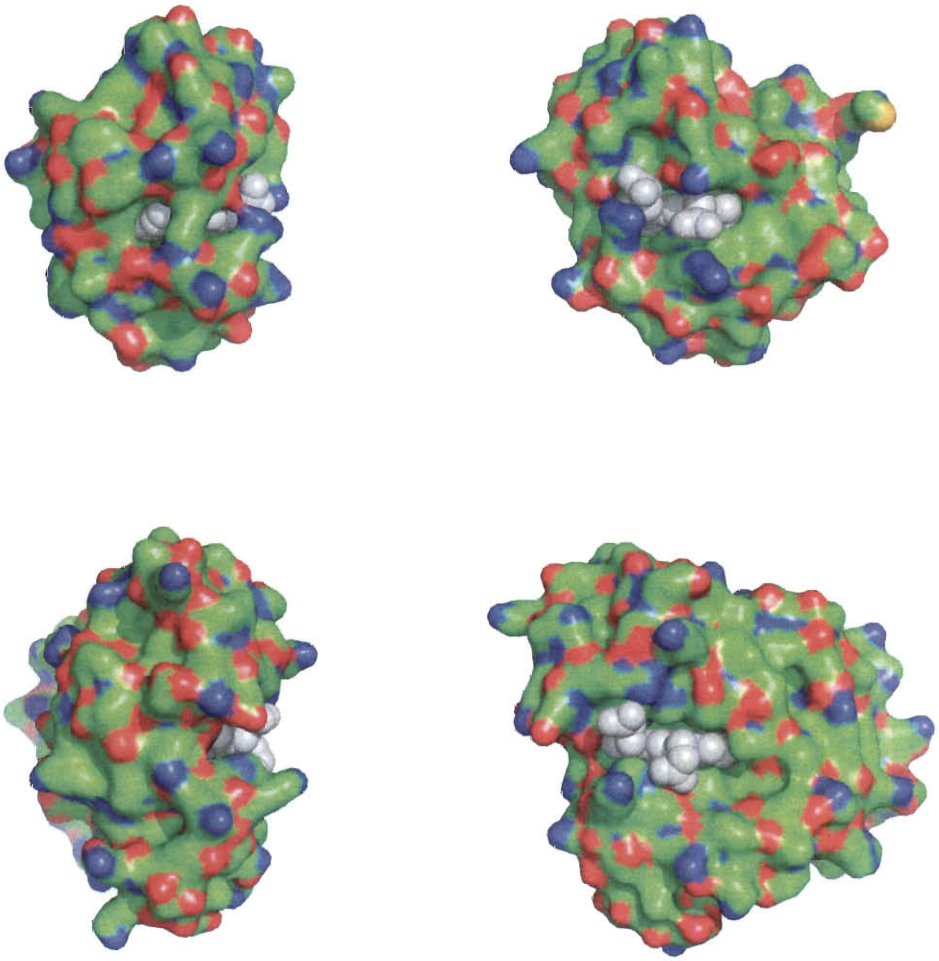


Figure 6.5. Space-filling representation of WT CerHb (upper panels) and WT sw Mb (lower panels). The heme is shown in light gray, hydrophobic residues in green, basic residues in blue and acidic residues in red. The left and right panels depict the same globin, but are rotated axially around the heme molecules. Images created using PyMol and the PDB entries: 1kr7 (CerHb) and 1jp6 (sw Mb).

6.5. Conclusions

Our results show that the apo-form of CerHb unfolds by a two-state mechanism and that a four-state unfolding model may be used to interpret the unfolding of the holo-protein (Fig. 6.2). The estimated value for K_{NH}^{0} is higher than expected, with the hemin affinity ~ 1300 fold weaker than that of WT sw Mb. Aquomet CerHb loses hemin at a rate comparable to that sw H97D Mb which is ≥ 500 times faster than hemin loss from sw WT Mb (Hargrove et al., 1996b). CerHb is significantly smaller than Mb and appears to have a less protected apolar pocket for hemin binding (Fig. 6.5), perhaps accounting for higher rate of hemin dissociation. The m_{NH} value obtained from the GuHCl dependence of rates of hemin loss is comparable to those for the metMb variants, which suggests that the *m*-values for the equilibrium hemin-dissociation constants reflect an inherent property of GuHCl interacting with free hemin and the corresponding apoglobins. In addition, the lack of a populated I or IH state is supported by the lack of any observable hemichrome formation. Thus, CerHb can serve as a simple globin model system for unfolding and assembly.

Chapter 7

The unfolding of human hemoglobin A

This chapter is adapted from:

Culbertson, D.S., and Olson, J.S. (2010). Folding and Assembly of Myoglobins and Hemoglobins. In Protein Folding and Metal Ions: Mechanisms, Biology, and Disease, P. Wittung-Stafshede, and C.M. Gomes, eds. (Taylor and Francis, Inc.) pp. 97-122.

7.1. Introduction

The exact assembly pathway for heterotetrameric HbA has remained unresolved. Possible mechanisms are explored in this Chapter starting with the scheme for the unfolding of the $\alpha_1\beta_1$ apoHb dimer, and followed by measurements of holo-metHb unfolding that show the appearance of hemichrome intermediate identical to those observed during holoMb denaturation.

7.2. Theoretical investigation of the folding mechanism for the $\alpha_1\beta_1$ apoHb dimer

Removal of heme from native HbA generates an $\alpha_1\beta_1$ apoHb dimer. The unfolding of this dimer into denatured monomeric subunits can occur by several mechanisms. The three most likely pathways for unfolding are shown in Fig. 7.1. For simplicity, the α and β subunits are assumed to have similar properties.

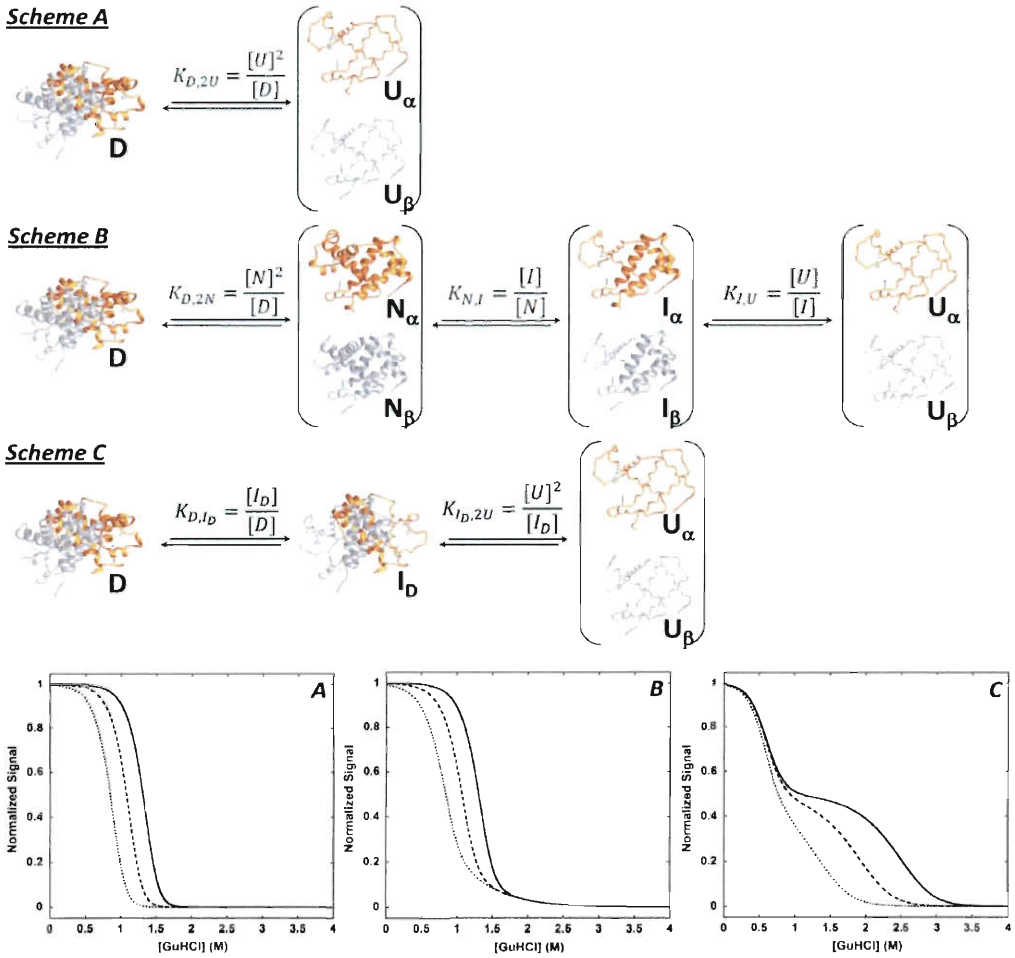


Figure 7.1. Equilibrium-unfolding mechanisms for the $\alpha_1\beta_1$ apoHb dimer. Possible unfolding mechanisms for apo-Hb dimers and simulations of GuHCl-induced equilibrium unfolding based on these mechanisms. D, I_D, N, I and U represent the dimer, dimer intermediate, native, intermediate and unfolded states. **Upper panel, Scheme A.** Model in which the apoHb dimer dissociates and unfolds in one step. $K_{2U,D} = 6.25 \times 10^{10}$ M^{-1}, $m_{2U,D} = 7.2$ kcal mol^{-1} M^{-1} **Scheme B.** Model in which the dimer unfolds to two folded states followed by partial unfolding to an intermediate I and then complete unfolding. $K_{2N,D} = 1.0 \times 10^8$ M^{-1}, $m_{2N,D} = 0.5$ kcal mol^{-1} M^{-1}, $K_{I,N} = 5$, $m_{I,N} = 2.2$ kcal mol^{-1} M^{-1}, $K_{U,I} = 5$, $m_{U,I} = 1.2$ kcal mol^{-1} M^{-1}. **Scheme C.** Model involving a partially folded dimer intermediate and then dissociation into unfolded monomers. $K_{ID,D} = 100$, $m_{ID,D} = 4.3$ kcal mol^{-1} M^{-1}, $K_{2U,ID} = 6.25 \times 10^8$ M^{-1} and $m_{2U,ID} = 2.9$ kcal mol^{-1} M^{-1}. The K and m-values in all three models were set to estimated values that, were mathematically equivalent and gave mid-point GuHCl concentrations similar to those observed in experiment by Graves et al (2008) at 5 °C, pH 7 for P_0 ~3-10 μM. The α and β subunits were created with PyMol and PDB structure 1ird and, for these mechanisms, are assumed to be equivalent in stability. **Lower panel,** Simulations for the unfolding of the apoHb dimer according to three models depicted in panels **A**, **B** and **C**, respectively. Protein concentration was varied from 1 (dotted) to 20 (dashed) to 400 (solid line) μM. Equations for all three models are given in the Appendices A.7-9. Note, the constants are folding, not unfolding constants.

In mechanism A, the dimer is assumed to unfold and dissociate into unfolded monomers in a single concerted process (Fig. 7.1A). In mechanism B, the dimer dissociates into folded monomers, which in turn unfold via a monomeric intermediate into completely unfolded monomers (as observed for apoMb, Fig. 7.1B). In mechanism C the dimer first unfolds to an intermediate with ~50% loss of helicity, but with the subunits still associated by the $\alpha_1\beta_1$ interface. Then, this dimeric I state dissociates into completely unfolded monomers (Fig. 7.1C). Because dissociation into monomers occurs in each case, all three apo-dimer mechanisms predict a strong dependence on total protein concentration, with increased resistance to unfolding at high P_0. Derivations of expressions for the unfolding curves predicted by each mechanism are given in the Appendix A.7-9. These equations were used to simulate GuHCl-induced unfolding curves as a function of total protein concentration, P_0, and results for 1, 20, and 400 μM apo-Hb subunits are shown in Fig. 7.1.

The unfolding curves predicted by mechanism A are steep, as expected for a simple two-state mechanism (*i.e.*, D or 2U), and show a marked right-shift to greater stability at higher protein concentrations. Although more complex with four states, mechanism B also predicts steep unfolding curves if reasonable values of the folding parameters are used. In this mechanism, the major resistance to unfolding is dimer dissociation into the much less stable monomers, which unfold completely at the denaturant concentrations required to disrupt the interface of the original apoHb dimer. Again, a strong dependence on protein concentration is predicted. In contrast, mechanism C predicts broad unfolding curves, which become biphasic at high protein concentrations. In this case, the initial step involves a unimolecular conformational change, which is

independent of protein concentration because the dimer remains intact. The second step in mechanism C is bimolecular involving dissociation into unstable monomers, which is markedly inhibited at high protein concentrations.

Preliminary data collected by Philip E. Graves (Graves, 2008) indicated that apoHb dimers show broad unfolding curves, which begin to show two distinct phases of roughly equal amplitudes at high apoHb concentrations. These data suggest that mechanism C applies to apoHbA. The simplest structural interpretation is that, like apoMb, initial addition of denaturant causes unfolding of the heme pocket and its associated B, C, D, and E helices, whereas the A, G, and H helices remain intact in both the α and β subunits, preserving most of the $\alpha_1\beta_1$ interface, which involves helices B, G, and H in the intact Hb tetramer. Further addition of denaturant simultaneously disrupts the dimer interface and unfolds the A, G, and H helices leading to completely unfolded subunits. Preliminary efforts were made to repeat these equilibrium-unfolding experiments with native apoHb dimers. These initial experiments gave results similar to Graves data, but the intrinsic unstability of apoHbA causes large problems with variability in the data at low [GuHCl] and precipitation at high protein concentrations. More work is needed to construct more stable HbA mutants that will allow more detailed experimental analyses. However, our theoretical analysis is new and novel, and likely to be applicable for future analyses.

7.3. Presence of hemichrome state in the unfolding of metHbA

The presence of bi-histidyl hemichrome adducts is well documented in hemoglobinopathies or thalassemias (Asakura et al., 1977; Bank et al., 1985; Kan, 1985;

Weiss et al., 2005; Winterbourn and Carrell, 1974). Hemichromes are notably observed in Hb aggregates of Heinz bodies (intra-erythrocytic aggregation of denatured Hb) as reviewed by Winterbourn (1990) and have been linked to denatured globins for over 50 years (Falbe-Hansen, 1961) (Fig. 7.2). These hemichromes were identified spectrally in α and β thalassemia erythrocytes and were considered to be Hb states that were irreversibly denatured (Asakura et al., 1977; Bunn and Forget, 1986).

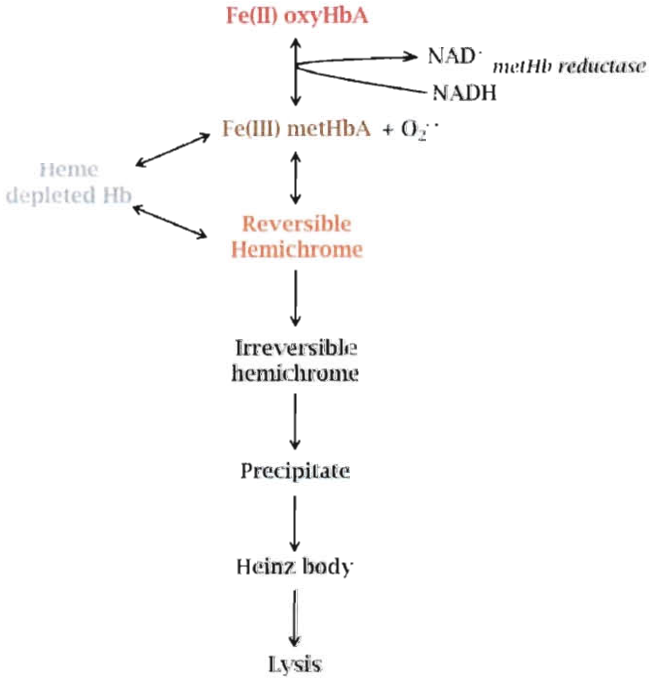


Figure 7.2. Degradation pathway for HbA. Adapted from (Winterbourn and Carrell, 1974).

More recently, hemichromes have been shown to be naturally occurring hexacoordinate states of Hbs from several vertebrate species including neuroglobins and cytoglobins (Kakar et al., 2010; Mitchell et al., 1995; Vergara et al., 2009; Vergara et al., 2007; Vergara et al., 2008). In adult humans, hemichromes are believed to be on the degradation pathway of HbA once oxidation has occurred (Arnold et al., 1999; Rachmilewitz et al., 1971; Winterbourn and Carrell, 1974). In Chapter 4, we have

provided evidence that GuHCl-induced hemichromes are reversible in sw Mb, and that they are on the pathway to globin assembly. In contrast, we were not able to "find" a hemichrome species during the unfolding of holoCerHb, which does not show an intermediate during unfolding of its apoglobin form. Thus, hemichrome formation appears to be associated with hemin binding to an intermediate (I) state with the heme pocket unfolded.

As a further test of these ideas and to search for intermediate folding states in HbA, we examined GuHCl-induced equilibrium unfolding of native aquomet HbA at 10 and 100 μM concentrations. As in previous chapters, unfolding was followed by CD, UV-visible absorbance, and in some cases, fluorescence emission changes.

The presence of a reversible hemichrome state during unfolding of metHbA is clearly evident from the visible-absorbance spectra. The Soret band red-shifts to 414-415 nm, and there is a strong appearance of the characteristic low spin bands at 535 and 565 nm (Fig. 7.3). Hemichrome species are strongly present even at concentrations of 10 μM HbA on a per heme basis, and increasing the metHb concentration to 100 μM increases population of the hemichrome species. The dimerization of hemin is also apparent at 100 μM metHb, with the appearance of a 370 nm band (Fig. 7.3).

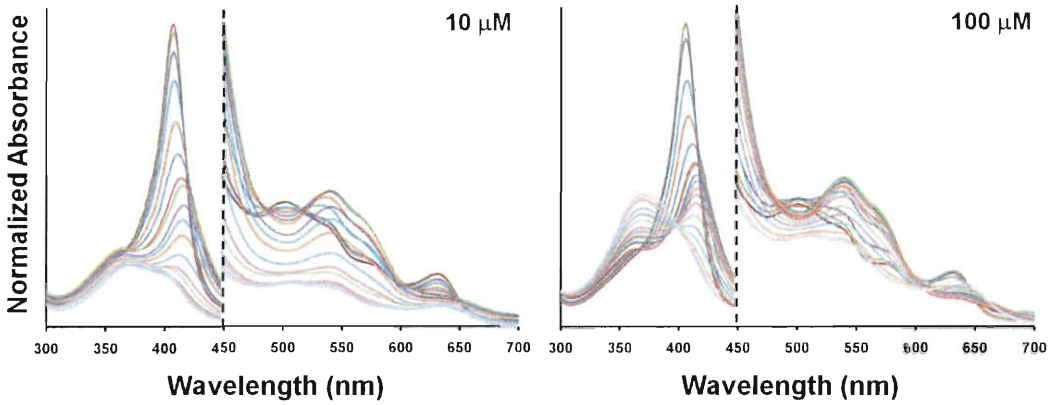


Figure 7.3. GuHCl-induced equilibrium-unfolding of native human aquometHbA monitored by UV-visible absorbance. The 10 and 100 μM of HbA concentrations are depicted in the left and right panels, respectively. The absorbance signals were normalized, and the absorbance scale from 450 to 700 nm is amplified 7-fold. The presence of the hemichrome is noticeable with the dual 535-565 nm band. Conditions: 100 mM potassium phosphate pH 7, 20 °C.

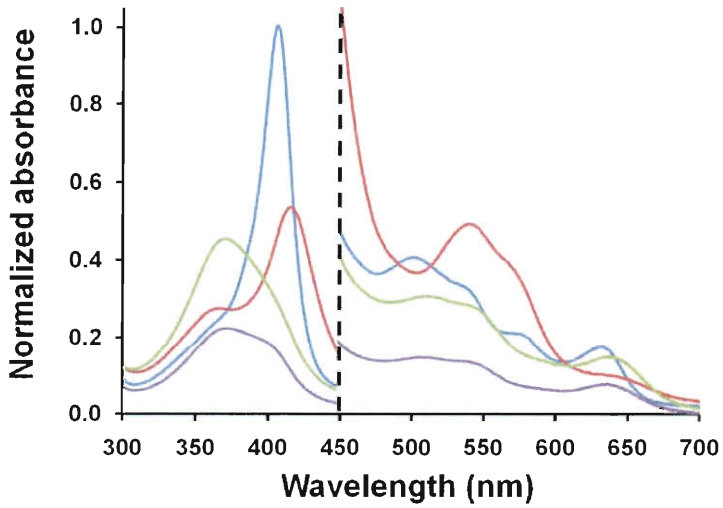


Figure 7.4. Deconvoluted absorbance spectra of native aquometHbA species present during equilibrium-unfolding. Three species are observed in the GuHCl titrations shown in Fig. 7.2: native (blue), hemichrome (red) and free hemin in 4 M GuHCl (green and purple). The spectra were normalized to 1 based on the native signal for aquometHbA. The purple and green curves depict the spectra for the unfolding endpoints at 10 and 100 μM globin concentration on a per heme basis, respectively. The 450 to 700 nm range is on a 0-0.15 absorbance scale.

Only three species appear to be responsible for the overall absorbance spectra in Fig. 7.3: native metHb, a hemichrome intermediate, and hemin (Fig. 7.4). The hemichrome species is readily observable with a Soret band at 416 nm, and the characteristic low spin 535-565 nm bands (red spectrum, Fig. 7.4). The spectra at 4 M GuHCl are different at high versus low protein concentration (green vs. purple spectra in Fig. 7.4), as observed by the increase in absorbance at 370 nm at 100 μM HbA.

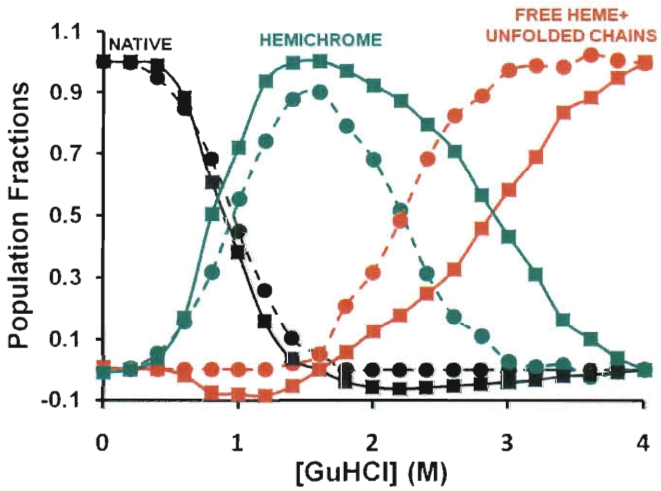


Figure 7.5. Deconvolution of the populations of the three-states present during equilibrium-unfolding of native aquometHbA. The unfolding was induced by addition of GuHCl and the process was followed by UV-vis absorbance. The deconvolutions were based on the equilibrium-unfolding depicted in Fig. 7.3 and on three observable states: native (black), hemichrome (green) and hemin with unfolded apoHb chains (red). The 10 and 100 μM concentrations are shown in filled circles and squares, with dashed and solid lines, respectively. Conditions: 100 mM potassium phosphate pH 7, 20 °C.

In an attempt to understand how HbA unfolds, I determined the fractional amount of each of the three species from the absorbance spectra at each GuHCl concentration using the standards in Fig. 7.4 for native metHbA, the hemichrome, and hemin. The fractions of each spectral species during unfolding are shown as a function of [GuHCl] in Fig. 7.5. The hemichrome species dominates at ~1.6 M GuHCl at both protein concentrations, but persists longer at the higher HbA concentration. Mechanistically, this

suggests that hemin dissociation occurs during unfolding of the hemichrome state, which should be inhibited at higher HbA concentration. The exact structure of the hemichrome and which histidines (and possible methionine) are coordinating the heme cofactors is still unknown as in sw Mb (see Chapter 4). Unfortunately, these data and analyses are too preliminary to derive detailed structure conclusions about the nature of the hemichrome species; however there is no doubt that it exists reversibly during metHbA unfolding induced by GuHCl.

7.4. Possible folding and assembly mechanism for holoHb

A possible mechanism for Hb assembly is shown in Fig. 7.6 and is based in part on the broad apoHb unfolding curves, observed in our initial experiments and simulated by mechanism C in Fig. 7.1. In this more complex scheme, heme binding only occurs after the folded apoHb dimer is formed. However, as discussed in Chapter 4, studies of holoMb unfolding suggest that heme can bind non-specifically to folding intermediates and possibly to the completely unfolded state, albeit with a lower affinity than to the N or I folded states. If heme binding occurs to all species involved in the assembly of tetrameric hemoglobin formation, the complex eight-state mechanism shown in Fig. 7.6 would apply.

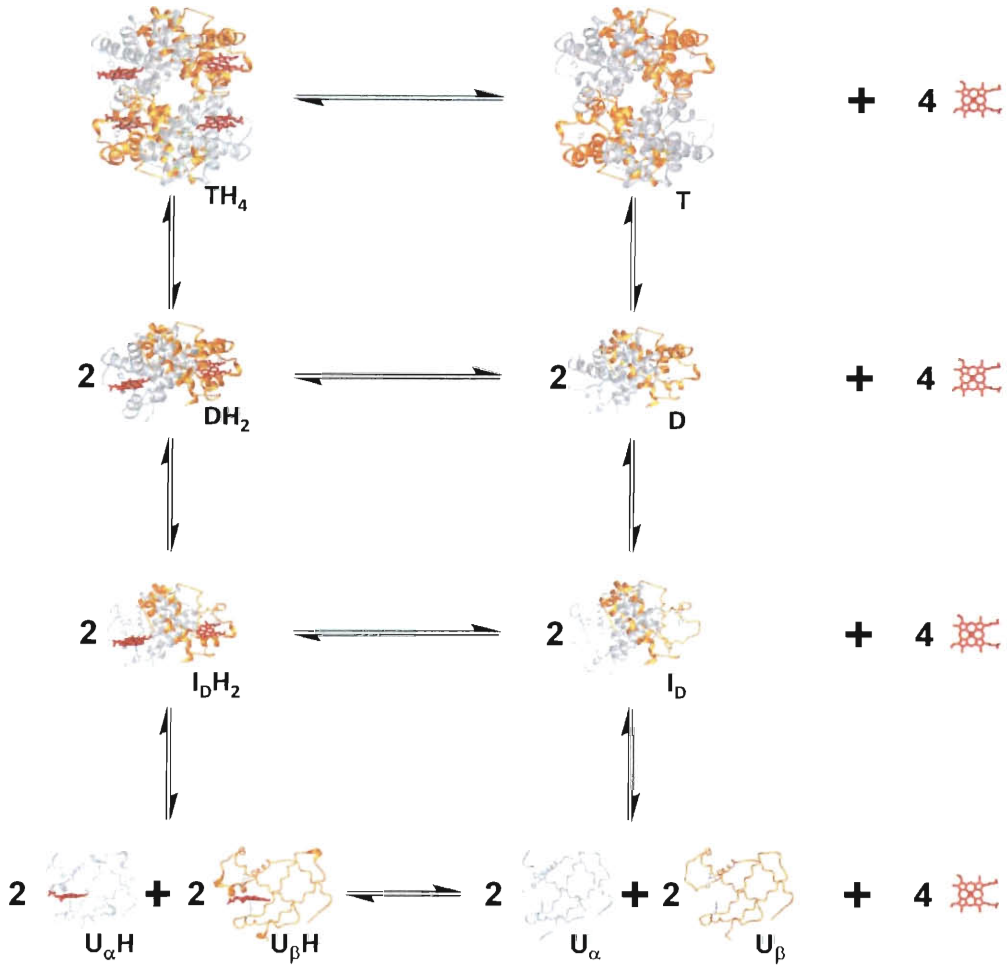


Figure 7.6. Proposed mechanistic model for the assembly of HbA. In this scheme, we have assumed that folded apoHb is present as an $\alpha_1\beta_1$ dimer, and that this dimer is formed via a dimer intermediate. We also assumed that the apoHbA dimer could potentially redimerize into a heterotetramer at high enough concentration, and that heme could bind to all apo states. A hemichrome state may be represented as the I_DH_2 species. The ribbon representations of the structures were created by Jayashree Soman using Ribbons and PDB entry: 1ird, for native holoHbA (TH_4). The other structures were based on the folding of apoMb and are only for illustrative purposes.

This mechanism requires a definition of the apoHb folding constants describing conversion of completely unfolded apo-subunits to apo-tetramers, plus heme binding to each of the four apoglobin states. Deriving the expressions for the folded species is much more difficult than for Mb, which only involves monomeric species, and requires finding roots of depressed quartic equations due to tetramer formation and heme dissociation, and

leads to a set of transcendental equations. The real situation may be even more complex if differences between the heme affinities and stabilities of the α and β subunits are large and if heme dimerization occurs. Thus, even the scheme in Fig. 7.6 is a simplification.

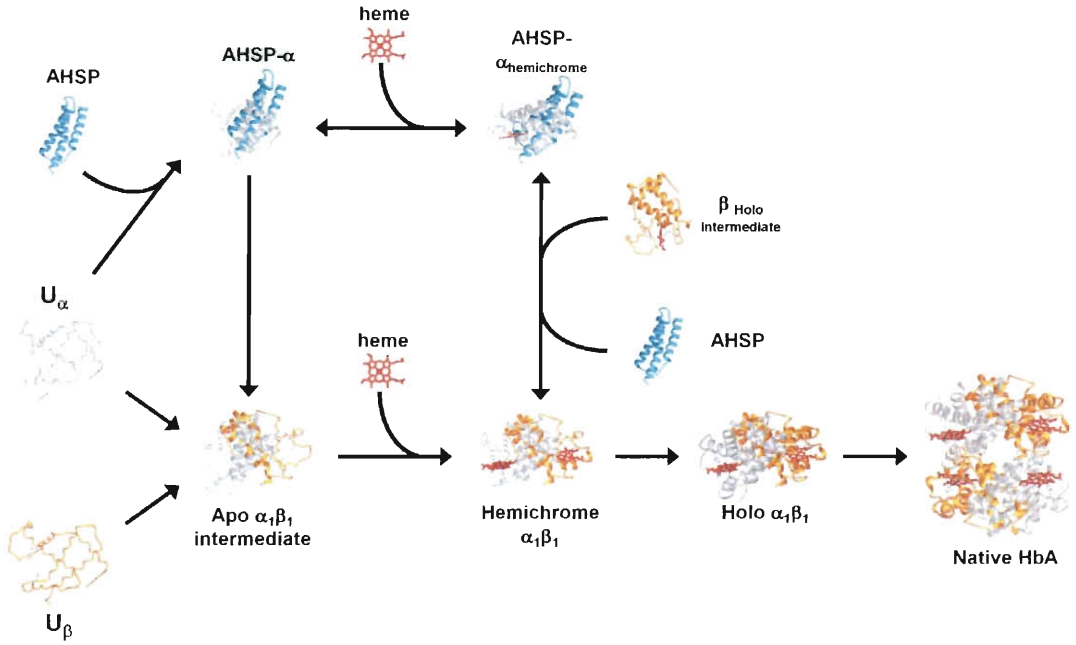


Figure 7.7. Proposed assembly mechanism for HbA. The mechanism was based on the assumptions that hemichrome species are dimeric globins as represented by I_DH_2 and that they are on path to assembly. Hemin would bind to the apo dimer intermediate, and in the cell reduction would induce folding and assembly to the native tetramer.

The presence of a reversible hemichrome HbA species, induced by the presence of GuHCl, is of great value to the understanding of HbA assembly. If heme stays in the ferric form due to the presence of oxygen in the cytosol, the hemichrome species as identified in Fig. 7.4 could be on-path to assembly. Upon binding of the hemin cofactors to the globin chains, the reducing environment could induce rapid folding and formation of oxy and deoxyHb (Fe(II)) tetramers (Fig. 7.7). The newly discovered chaperone, alpha hemoglobin stabilizing protein (AHSP), aids in the assembly of Hb in the presence of

excess α-chains (Kihm et al., 2002) and upon binding to the holo-α-chains, induces a hemichrome conformation (Feng et al., 2004; Feng et al., 2005), reinforcing the idea that hemichrome species are on-path to assembly. However, the addition of AHSP adds another level of complexity to the assembly mechanism for HbA and would include folding steps for the α subunits alone. The assembly mechanism presented in Fig. 7.6 will require testing in future more quantitative analyses.

Chapter 8

Conclusions

In the past, the role of heme in the folding and assembly of globins remained elusive because of the lack of a quantitative theory to analyze observed unfolding curves. Most hemoglobin and myoglobin studies focused on the apoprotein without investigating how heme modulates the stability of the hologlobin. In this thesis, an analytical approach for interpreting the equilibrium-unfolding of monomeric hologlobins was derived and used to estimate the equilibrium constants for the dissociation of hemin from the globin folding states. This method requires systematic measurements of the unfolding of both apo- and holo-forms. To our knowledge, this combined analysis had never been done before, and the results in Chapters 4, 5, and 6 provide strong evidence for the utility of this methodology for determining quantatively the effect of hemin binding on Mb stability. In theory, this approach is not limited to globins but applies to the study of all cofactor-containing proteins.

The six-state mechanism presented in Chapter 4 (Fig. 4.3) was used to interpret the equilibrium-unfolding of five sw Mb variants and demonstrate that hemin binds specifically to the N apo state with very high affinity and to the apo molten globule intermediate (I) with moderate affinity ($\sim 10^{-11}$ M). The discovery of the latter GuHCl-induced hemichrome species, IH, during equilibrium-unfolding was unexpected. Hemichrome species have been the subject of great interest recently because hemichromes are present in the native states of several newly discovered hemoglobins, as

reviewed by Kakar et al. (2010). The hemichrome species in Mb is an equilibrium-unfolding intermediate state with hemin coordinated most likely by two strong-field ligand side chains, including the proximal histidine (His93) and either the distal histidine (His64) or, in its absence, a nearby methionine (Met55) or histidine side chain. Another unexpected result is that GuHCl markedly induces hemin dissociation, increasing K_{NH}, K_{IH}, and K_{UH}, and left-shifting the unfolding curves of the hologlobins.

In Chapter 5, I was able to analyze semi-quantatively the effects of hemin dimerization and protein concentration on GuHCl-induced equilibrium-unfolding of WT and H97D sw holoMb. The addition of a free hemin dimerization step significantly increased the complexity of the analyses. As a result, fitting the unfolding curves at three different concentrations simultaneously was far more difficult than anticipated due to the large number of parameters needed to simulate the curves, and several of these parameters were undefined by the observed data. The reason for the lack of an observable protein concentration dependence of the GuHCl unfolding midpoint has been the subject of debate. My results suggest that the relatively strong binding of hemin to the I state to form the IH intermediate is the main factor. Our conclusions are built on the fact that the unfolding curves only show a strong protein concentration dependence during the later stage of unfolding, when hemin dissociation from the I state occurs, and reattachment to the unfolded U state is unlikely at these high [GuHCl]. The affinity of unfolded Mb for hemin in buffer is on the order of 1×10^{-6} M when defined as a K_{UH}^{0} constant, and the presence of GuHCl weakens the affinity significantly as is observed for hemin binding to the N and I apoMb states. Contrary to our preliminary results with free hemin, our model suggests that GuHCl actually decreases free hemin dimerization at very high

concentrations of denaturant. Thus there is a build up and decay of this species, particularly at high total initial holoMb concentration.

We also tested our analytical expressions with CerHb, a mini-globin which is lacking the A helix and does not appear to populate a folding intermediate in the apo form. The lack of an observable hemichrome species in CerHb supports our hypothesis that WT apoCerHb unfolds by a two-state mechanism and does not populate a folding intermediate. This result suggests that hemichrome formation in Mb requires a molten globule intermediate state in which the hemin pocket is partially or completely unfolded.

In general, the level of recombinant expression of monomeric globins in *E. coli* appears to be correlated with the stability in the apo-form and hemin affinity. Thus it was initially surprising that CerHb expresses readily in *E. coli* despite low hemin affinity and stability in the folded apo-N state. However, unlike apoMb and particularly apoHb, the completely unfolded states of WT apoCerHb do not aggregate or precipitate rapidly at room temperature. Thus, even though apoCerHb unfolds more readily to the U state, it remains in solution, can readily refold in the presence of the cofactor, and then becomes more stable when the heme-iron is reduced. The cause of the increased solubility of the U state could be due to the smaller size of CerHb and its pI and charge distribution.

In an effort to demonstrate the physiological relevance of our biophysical approach, I have correlated the GuHCl-induced unfolding parameters with the thermal stabilities of the proteins as measured empirically by their melting temperatures. Melting-curves for apo- and hologlobin variants were monitored by the changes in ellipticity at 222 nm as an indicator of secondary structure content. In contrast to GuHCl-induced unfolding, thermal unfolding of both apo- and hologlobins is not reversible; however,

measuring the midpoint of the thermal-unfolding transition still provides a good indicator of relative thermal stability when comparing variants. As hoped, the marked differences in resistance of the apoglobin variants to unfolding by GuHCl correlate well with changes in thermal stability, which also varied widely in both the apo- and holo-forms (Table 8.1). The correlations shown in panel B of Fig. 8.1 also demonstrate that the GuHCl midpoints for unfolding of the apoMb variants correlate well with the Gibbs free energy calculated using parameters derived from fits of the observed data for Mb variants and CerHb.

Table 8.1. Comparison of thermal and chemical stability of apo- and hologlobins.

Variant	Apo		Holo	
	T_m (°C)	$[GuHCl]_{1/2}$ (M)	T_m (°C)	$[GuHCl]_{1/2}$ (M)
WT Mb	63.9	1.35	84.6	1.77
H64F Mb	72.0	1.60	85.0	1.93
T67P Mb	48.9	0.78	74.7	1.05
V68T Mb	56.6	1.02	87.7	2.03
H97D Mb	65.7	1.25	72.5	1.20
H64Y/V68F Mb	72.5	n.d.	n.d.	n.d.
WT CerHb	34.3	0.48[1]	66.4	0.82[1]
Native HbA	<20	0.95[2]	61.2	1.20[3]

The thermal midpoints were retrieved by fitting unfolding data induced by heat, monitored by CD at 222 nm and using a two-state thermal-unfolding equation (Sanchez-Ruiz et al., 1988). The GuHCl midpoints, $[GuHCl]_{1/2}$, were obtained at half the CD signal at 222 nm for the holo variants, and at maximal fluorescence intensity for the apoMb variants.

[1] The $[GuHC]_{1/2}$ for apo and holoCerHb were estimated using CD signals at 222 nm at 20°C in 100 mM phosphate pH 7.

[2] The $[GuHC]_{1/2}$ for apoHbA was estimated using CD signals at 222 nm, at 40 μM apoHb concentration (on a subunit basis) in 100 mM phosphate pH 7 at 4 °C.

[3] The $[GuHC]_{1/2}$ for native metHbA was estimated using CD signals at 222 nm, at 10 μM globin concentration (on a subunit basis), in 100 mM phosphate pH 7 at 20 °C.

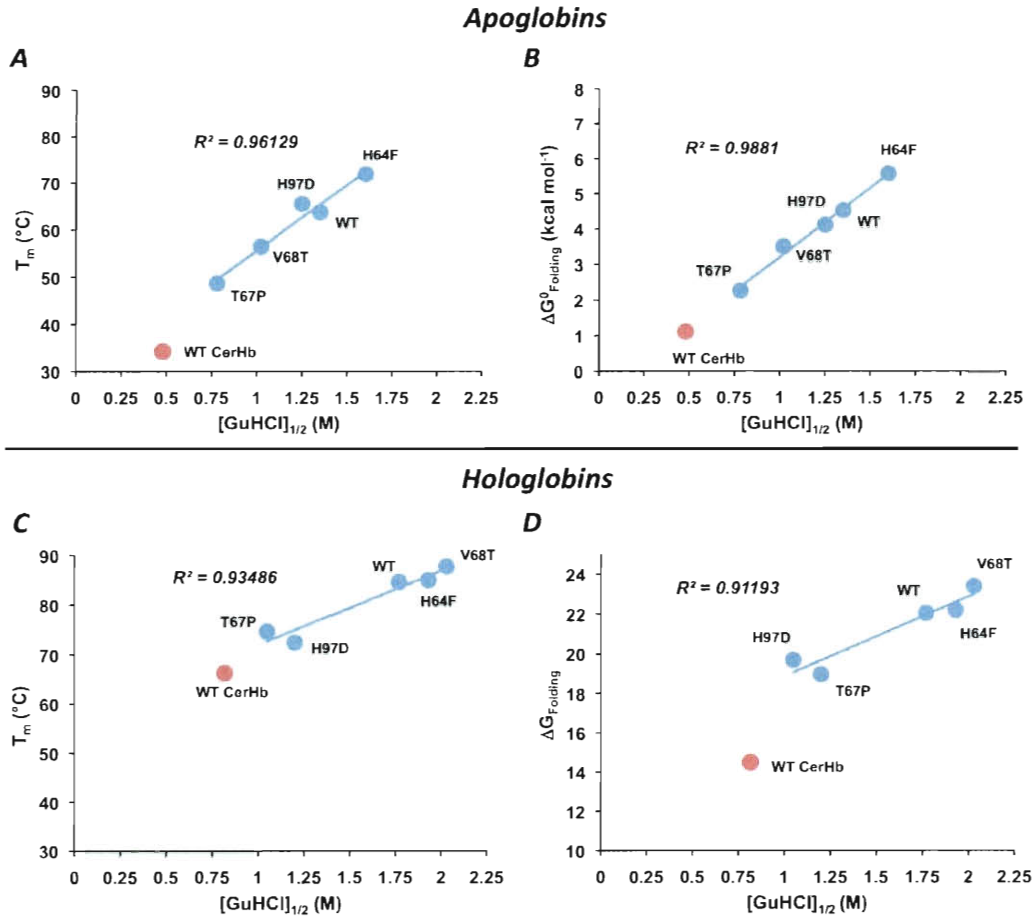


Figure 8.1. Correlations between thermal and chemical stability of apo and hologlobin variants. The five Mb variants from Chapter 4 are represented with WT CerHb from Chapter 6. CerHb is only shown for illustrative purposes as the unfolding was done at lower temperature. The thermal stability is expressed as thermal midpoint, T_m, and uses values from Table 8.1. The chemical stabilities are defined as the free energies of GuHCl-induced equilibrium unfolding (A and C), and as [GuHCl] transition midpoints (B and D). The values used are defined in Tables 8.1, 4.2, and 4.4. $\Delta G^0_{Folding}$ for the apoproteins was computed as $-RT\ln(K_{NI}^0 K_{IU}^0)$ and $\Delta G_{Folding}$ was computed as $-RT\ln(K_{NH}^0 K_{NI}^0 K_{IU}^0)$.

The correlations for the holoMb variants and CerHb in panels C and D of Fig. 8.1 also demonstrate that the resistance of these hologlobins can be expressed quantatively in terms of folding free energy (*i.e.*, $-RT\ln(K_{NH}^0 K_{NI}^0 K_{IU}^0)$) derived from our mechanism and that the relative values represent the observed orders of stabilities measured empirically by midpoint [GuHCl] values.

HbA populates an $\alpha_1\beta_1$ dimer upon extraction of heme, which is on the path to assembly of HbA. Using a relatively simple theory, I evaluated three possible mechanisms for the formation of the apoHbA dimer. Preliminary data suggest that the apo-dimer folds via an apo-dimer intermediate; however we were not able to confirm our hypothesis due to experimental difficulties caused by the low stability and propensity for aggregation of apoHbA. Future analyses with more stable apoHb mutants should resolve the exact mechanism by which the $\alpha_1\beta_1$ apoHbA dimer assembles.

It has been well documented that hemichrome species form during the denaturation of native HbA. Some of these hemichrome forms are found *in vivo* as red cell Heinz bodies. My results demonstrate that reversible hemichromes are also formed during equilibrium-unfolding of metHbA, and therefore are probably also on path for holoHbA assembly, which is much more complex. As a start, I have suggested an eight-state mechanism for the equilibrium-unfolding of HbA (Fig. 7.6), in which Hb assembles via an apo-dimer intermediate that binds hemin to form a hemichrome species designated I_DH_2 state. The spectral deconvolutions shown in Figs. 7.3-4 support this model and demonstrate that hemichromes are readily observable during holoHbA unfolding, and protein concentration dependence only occurs for the unfolding of hemichrome dimer species but not for that of the native holo-dimer state. Further studies are required to confirm our initial interpretations and to determine step by step the exact assembly mechanism for HbA as it occurs *in vivo*, with the addition of AHSP as a chaperone.

REFERENCES

Adams, P.A. (1976). The kinetics and mechanism of the recombination reaction between apomyoglobin and haemin. Biochem J *159*, 371-376.

Adams, P.A. (1977). The kinetics of the recombination reaction between apomyoglobin and alkaline haematin. Biochem J *163*, 153-158.

Anfinsen, C.B. (1972). The formation and stabilization of protein structure. Biochem J *128*, 737-749.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

Antonini, E., and Brunori, M. (1971). Hemoglobin and myoglobin in their reactions with ligands (Amsterdam, North-Holland Pub. Co.).

Antonini, E., and Gibson, Q.H. (1960). Some observations on the kinetics of the reactions with gases of natural and reconstituted haemoglobins. Biochem J *76*, 534-538.

Apiyo, D., and Wittung-Stafshede, P. (2002). Presence of the cofactor speeds up folding of Desulfovibrio desulfuricans flavodoxin. Protein Sci *11*, 1129-1135.

Appleby, C.A. (1984). Leghemoglobin and Rhizobium Respiration. Ann Rev Plant Physiol *35*, 443-478.

Arnold, E.V., Bohle, D.S., and Jordan, P.A. (1999). Reversible and irreversible hemichrome generation by the oxygenation of nitrosylmyoglobin. Biochemistry *38*, 4750-4756.

Arredondo-Peter, R., Hargrove, M.S., Sarath, G., Moran, J.F., Lohrman, J., Olson, J.S., and Klucas, R.V. (1997). Rice hemoglobins. Gene cloning, analysis, and O2-binding kinetics of a recombinant protein synthesized in Escherichia coli. Plant Physiol *115*, 1259-1266.

Asakura, T., Minakata, K., Adachi, K., Russell, M.O., and Schwartz, E. (1977). Denatured hemoglobin in sickle erythrocytes. J Clin Invest *59*, 633-640.

Ascoli, F., Fanelli, M.R., and Antonini, E. (1981). Preparation and properties of apohemoglobin and reconstituted hemoglobins. Methods Enzymol *76*, 72-87.

Asher, C., de Villiers, K.A., and Egan, T.J. (2009). Speciation of ferriprotoporphyrin IX in aqueous and mixed aqueous solution is controlled by solvent identity, pH, and salt concentration. Inorg Chem *48*, 7994-8003.

Auton, M., Holthauzen, L.M., and Bolen, D.W. (2007). Anatomy of energetic changes accompanying urea-induced protein denaturation. Proc Natl Acad Sci U S A *104*, 15317-15322.

Baglioni, C. (1968). Chromosomal and cytoplasmic regulation of haemoglobin synthesis. Bibl Haematol *29*, 1056-1063.

Baglioni, C., and Campana, T. (1967). Alpha-chain and globin: intermediates in the synthesis of rabbit hemoglobin. Eur J Biochem *2*, 480-492.

Baldwin, R.L. (1996). How Hofmeister ion interactions affect protein stability. Biophys J *71*, 2056-2063.

Balestrieri, C., Colonna, G., Giovane, A., Irace, G., and Servillo, L. (1976). Equilibrium evidence of non-single step transition during guanidine unfolding of apomyoglobins. FEBS Lett *66*, 60-64.

Ballew, R.M., Sabelko, J., and Gruebele, M. (1996). Direct observation of fast protein folding: the initial collapse of apomyoglobin. Proc Natl Acad Sci U S A *93*, 5759-5764.

Bank, A. (2007). AHSP: a novel hemoglobin helper. J Clin Invest *117*, 1746-1749.

Bank, A., Dobkin, C., Donovan-Peluso, M., and Young, K. (1985). Abnormal globin gene structure and expression in beta-thalassemia. Ann N Y Acad Sci *445*, 1-9.

Barrick, D., and Baldwin, R.L. (1993a). Stein and Moore Award address. The molten globule intermediate of apomyoglobin and the process of protein folding. Protein Sci *2*, 869-876.

Barrick, D., and Baldwin, R.L. (1993b). Three-state analysis of sperm whale apomyoglobin folding. Biochemistry *32*, 3790-3796.

Barrick, D., Hughson, F.M., and Baldwin, R.L. (1994). Molecular mechanisms of acid denaturation. The role of histidine residues in the partial unfolding of apomyoglobin. J Mol Biol *237*, 588-601.

Baryshnikova, E.N., Melnik, B.S., Finkelstein, A.V., Semisotnov, G.V., and Bychkova, V.E. (2005). Three-state protein folding: experimental determination of free-energy profile. Protein Sci *14*, 2658-2667.

Bellomo, R., Daskalakis, M., Parkin, G., and Boyce, N. (1991). Myoglobin clearance during acute continuous hemodiafiltration. Intensive Care Med *17*, 509.

Berova, N., Nakanishi, K., and Woody, R.W. (2000). Circular Dichroism: principles and applications, 2nd Edition edn (Wiley-VCH).

Birnbaum, G.I., Evans, S.V., Przybylska, M., and Rose, D.R. (1994). 1.70 A resolution structure of myoglobin from yellowfin tuna. An example of a myoglobin lacking the D helix. Acta Crystallogr D Biol Crystallogr *50*, 283-289.

Braun, S.R., Weiss, F.R., Keller, A.I., Ciccone, J.R., and Preuss, H.G. (1970). Evaluation of the renal toxicity of heme proteins and their derivatives: a role in the genesis of acute tubule necrosis. J Exp Med *131*, 443-460.

Brittain, T., Skommer, J., Raychaudhuri, S., and Birch, N. (2010). An antiapoptotic neuroprotective role for neuroglobin. Int J Mol Sci *11*, 2306-2321.

Brown, G.C. (2003). Cell biology. NO says yes to mitochondria. Science *299*, 838-839.

Brown, S.B., Dean, T.C., and Jones, P. (1970). Aggregation of ferrihaems. Dimerization and protolytic equilibria of protoferrihaem and deuteroferrihaem in aqueous solution. Biochem J *117*, 733-739.

Bunn, H.F. (1987). Subunit assembly of hemoglobin: an important determinant of hematologic phenotype. Blood *69*, 1-6.

Bunn, H.F.a., and Forget, B.G. (1986). Hemoglobin: Molecular, Genetic and Clinical Aspects, 2 edn (Philadelphia, Elsevier Health Sciences).

Burmester, T., Gerlach, F., and Hankeln, T. (2007). Regulation and role of neuroglobin and cytoglobin under hypoxia. Adv Exp Med Biol *618*, 169-180.

Burmester, T., Haberkamp, M., Mitz, S., Roesner, A., Schmidt, M., Ebner, B., Gerlach, F., Fuchs, C., and Hankeln, T. (2004). Neuroglobin and cytoglobin: genes, proteins and evolution. IUBMB Life *56*, 703-707.

Carver, T.E., Brantley, R.E., Jr., Singleton, E.W., Arduini, R.M., Quillin, M.L., Phillips, G.N., Jr., and Olson, J.S. (1992). A novel site-directed mutant of myoglobin with an unusually high O2 affinity and low autooxidation rate. J Biol Chem *267*, 14443-14450.

Cassoly, R., and Banerjee, R. (1971). Structure and function of human semihemoglobins alpha and beta. Eur J Biochem *19*, 514-522.

Chatzizisis, Y.S., Misirli, G., Hatzitolios, A.I., and Giannoglou, G.D. (2008). The syndrome of rhabdomyolysis: complications and treatment. Eur J Intern Med *19*, 568-574.

Chen, J.J. (2007). Regulation of protein synthesis by the heme-regulated eIF2alpha kinase: relevance to anemias. Blood *109*, 2693-2699.

Chiba, K., Ikai, A., Kawamura-Konishi, Y., and Kihara, H. (1994). Kinetic study on myoglobin refolding monitored by five optical probe stopped-flow methods. Proteins *19*, 110-119.

Cocco, M.J., and Lecomte, J.T. (1990). Characterization of hydrophobic cores in apomyoglobin: a proton NMR spectroscopy study. Biochemistry *29*, 11067-11072.

Crespin, M.O., Boys, B.L., and Konermann, L. (2005). The reconstitution of unfolded myoglobin with hemin dicyanide is not accelerated by fly-casting. FEBS Lett *579*, 271-274.

Culbertson, D.S., and Olson, J.S. (2010). Role of heme in the unfolding and assembly of myoglobin. Biochemistry *49*, 6052-6063.

Daher Ede, F., Silva Junior, G.B., Brunetta, D.M., Pontes, L.B., and Bezerra, G.P. (2005). Rhabdomyolysis and acute renal failure after strenuous exercise and alcohol abuse: case report and literature review. Sao Paulo Med J *123*, 33-37.

de Villiers, K.A., Kaschula, C.H., Egan, T.J., and Marques, H.M. (2007). Speciation and structure of ferriprotoporphyrin IX in aqueous solution: spectroscopic and diffusion measurements demonstrate dimerization, but not mu-oxo dimer formation. J Biol Inorg Chem *12*, 101-117.

Dickerson, R.E., and Geis, I. (1983). Hemoglobin : structure, function, evolution, and pathology (Benjamin/Cummings Pub. Co.).

Dintzis, H.M. (1961). Assembly of the peptide chains of hemoglobin. Proc Natl Acad Sci U S A *47*, 247-261.

Dobson, C.M. (2002). Protein misfolding and its link with human disease. Paper presented at: Molecular Informatics: Confronting Complexity (Bozen, Italy).

Doherty, D.H., Doyle, M.P., Curry, S.R., Vali, R.J., Fattor, T.J., Olson, J.S., and Lemon, D.D. (1998). Rate of reaction with nitric oxide determines the hypertensive effect of cell-free hemoglobin. Nat Biotechnol *16*, 672-676.

Dordas, C. (2009). Nonsymbiotic hemoglobins and stress tolerance in plants. Plant Science *176*, 433-440.

Dordas, C., Rivoal, J., and Hill, R.D. (2003). Plant haemoglobins, nitric oxide and hypoxic stress. Ann Bot (Lond) *91 Spec No*, 173-178.

Dou, Y., Maillett, D.H., Eich, R.F., and Olson, J.S. (2002). Myoglobin as a model system for designing heme protein based blood substitutes. Biophys Chem *98*, 127-148.

Eakin, C.M., Knight, J.D., Morgan, C.J., Gelfand, M.A., and Miranker, A.D. (2002). Formation of a copper specific binding site in non-native states of beta-2-microglobulin. Biochemistry *41*, 10646-10656.

Ebert, B.L., and Bunn, H.F. (1999). Regulation of the erythropoietin gene. Blood *94*, 1864-1877.

Eich, R.F., Li, T., Lemon, D.D., Doherty, D.H., Curry, S.R., Aitken, J.F., Mathews, A.J., Johnson, K.A., Smith, R.D., Phillips, G.N., Jr., *et al.* (1996). Mechanism of NO-induced oxidation of myoglobin and hemoglobin. Biochemistry *35*, 6976-6983.

Eliezer, D., Chung, J., Dyson, H.J., and Wright, P.E. (2000). Native and non-native secondary structure and dynamics in the pH 4 intermediate of apomyoglobin. Biochemistry *39*, 2894-2901.

Eliezer, D., and Wright, P.E. (1996). Is apomyoglobin a molten globule? Structural characterization by NMR. J Mol Biol *263*, 531-538.

Eliezer, D., Yao, J., Dyson, H.J., and Wright, P.E. (1998). Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. Nat Struct Biol *5*, 148-155.

Ervin, J., Larios, E., Osvath, S., Schulten, K., and Gruebele, M. (2002). What causes hyperfluorescence: folding intermediates or conformationally flexible native states? Biophys J *83*, 473-483.

Evans, S.V., and Brayer, G.D. (1990). High-resolution study of the three-dimensional structure of horse heart metmyoglobin. J Mol Biol *213*, 885-897.

Falbe-Hansen, I. (1961). Spectrophotometric studies on denatured globin hemichrome. Scand J Clin Lab Invest *13*, 429-433.

Fandrich, M., Fletcher, M.A., and Dobson, C.M. (2001). Amyloid fibrils from muscle myoglobin. Nature *410*, 165-166.

Fanelli, A.R., Antonini, E., and Caputo, A. (1958). Studies on the structure of hemoglobin. I. Physicochemical properties of human globin. Biochim Biophys Acta *30*, 608-615.

Feng, L., Gell, D.A., Zhou, S., Gu, L., Kong, Y., Li, J., Hu, M., Yan, N., Lee, C., Rich, A.M., *et al.* (2004). Molecular mechanism of AHSP-mediated stabilization of alpha-hemoglobin. Cell *119*, 629-640.

Feng, L., Zhou, S., Gu, L., Gell, D.A., Mackay, J.P., Weiss, M.J., Gow, A.J., and Shi, Y. (2005). Structure of oxidized alpha-haemoglobin bound to AHSP reveals a protective mechanism for haem. Nature *435*, 697-701.

Ferreira, R., Ohneda, K., Yamamoto, M., and Philipsen, S. (2005). GATA1 function, a paradigm for transcription factors in hematopoiesis. Mol Cell Biol *25*, 1215-1227.

Flogel, U., Merx, M.W., Godecke, A., Decking, U.K., and Schrader, J. (2001). Myoglobin: A scavenger of bioactive NO. Proc Natl Acad Sci U S A *98*, 735-740.

Fordel, E., Geuens, E., Dewilde, S., De Coen, W., and Moens, L. (2004). Hypoxia/ischemia and the regulation of neuroglobin and cytoglobin expression. IUBMB Life *56*, 681-687.

Fordel, E., Thijs, L., Moens, L., and Dewilde, S. (2007). Neuroglobin and cytoglobin expression in mice. Evidence for a correlation with reactive oxygen species scavenging. Febs J *274*, 1312-1317.

Ganz, T. (2006). Hepcidin and its role in regulating systemic iron metabolism. Hematology Am Soc Hematol Educ Program, 29-35, 507.

Gardner, P.R., Gardner, A.M., Martin, L.A., Dou, Y., Li, T., Olson, J.S., Zhu, H., and Riggs, A.F. (2000). Nitric-oxide dioxygenase activity and function of flavohemoglobins. sensitivity to nitric oxide and carbon monoxide inhibition. J Biol Chem *275*, 31581-31587.

Garry, D.J., Ordway, G.A., Lorenz, J.N., Radford, N.B., Chin, E.R., Grange, R.W., Bassel-Duby, R., and Williams, R.S. (1998). Mice without myoglobin. Nature *395*, 905-908.

Gell, D., Kong, Y., Eaton, S.A., Weiss, M.J., and Mackay, J.P. (2002). Biophysical characterization of the alpha-globin binding protein alpha-hemoglobin stabilizing protein. J Biol Chem *277*, 40602-40609.

Giangiacomo, L., Ilari, A., Boffi, A., Morea, V., and Chiancone, E. (2005). The truncated oxygen-avid hemoglobin from Bacillus subtilis: X-ray structure and ligand binding properties. J Biol Chem *280*, 9192-9202.

Gibson, Q.H., and Antonini, E. (1960). Kinetic studies on the reaction between native globin and haem derivatives. Biochem J *77*, 328-341.

Goustin, A.S., and Wilt, F.H. (1982). Direct measurement of histone peptide elongation rate in cleaving sea urchin embryos. Biochim Biophys Acta *699*, 22-27.

Graca-Souza, A.V., Arruda, M.A., de Freitas, M.S., Barja-Fidalgo, C., and Oliveira, P.L. (2002). Neutrophil activation by heme: implications for inflammatory processes. Blood *99*, 4160-4165.

Graves, P.E. (2008). Enhancing Stability and Expression of Recombinant Human Hemoglobin in E. coli: Progress in the Development of a Recombinant HBOC Source (Rice University).

Graves, P.E., Henderson, D.P., Horstman, M.J., Solomon, B.J., and Olson, J.S. (2008). Enhancing stability and expression of recombinant human hemoglobin in E. coli: Progress in the development of a recombinant HBOC source. Biochim Biophys Acta *1784*, 1471-1479.

Gray, H.B. (2003). Biological inorganic chemistry at the beginning of the 21st century. Proc Natl Acad Sci U S A *100*, 3563-3568.

Griko, Y.V., Privalov, P.L., Venyaminov, S.Y., and Kutyshenko, V.P. (1988). Thermodynamic study of the apomyoglobin structure. J Mol Biol *202*, 127-138.

Hamza, I. (2006). Intracellular trafficking of porphyrins. ACS Chem Biol *1*, 627-629.

Han, A.P., Yu, C., Lu, L., Fujiwara, Y., Browne, C., Chin, G., Fleming, M., Leboulch, P., Orkin, S.H., and Chen, J.J. (2001). Heme-regulated eIF2alpha kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency. Embo J *20*, 6909-6918.

Hankeln, T., Ebner, B., Fuchs, C., Gerlach, F., Haberkamp, M., Laufs, T.L., Roesner, A., Schmidt, M., Weich, B., Wystub, S., *et al.* (2005). Neuroglobin and cytoglobin in search of their role in the vertebrate globin family. J Inorg Biochem *99*, 110-119.

Hargrove, M.S., Barrick, D., and Olson, J.S. (1996a). The association rate constant for heme binding to globin is independent of protein structure. Biochemistry *35*, 11293-11299.

Hargrove, M.S., Krzywda, S., Wilkinson, A.J., Dou, Y., Ikeda-Saito, M., and Olson, J.S. (1994a). Stability of myoglobin: a model for the folding of heme proteins. Biochemistry *33*, 11767-11775.

Hargrove, M.S., and Olson, J.S. (1996). The stability of holomyoglobin is determined by heme affinity. Biochemistry *35*, 11310-11318.

Hargrove, M.S., Singleton, E.W., Quillin, M.L., Ortiz, L.A., Phillips, G.N., Jr., Olson, J.S., and Mathews, A.J. (1994b). His64(E7)-->Tyr apomyoglobin as a reagent for measuring rates of hemin dissociation. J Biol Chem *269*, 4207-4214.

Hargrove, M.S., Whitaker, T., Olson, J.S., Vali, R.J., and Mathews, A.J. (1997). Quaternary structure regulates hemin dissociation from human hemoglobin. J Biol Chem *272*, 17385-17389.

Hargrove, M.S., Wilkinson, A.J., and Olson, J.S. (1996b). Structural factors governing hemin dissociation from metmyoglobin. Biochemistry *35*, 11300-11309.

Hemmaplardh, D., and Morgan, E.H. (1974a). The mechanism of iron exchange between synthetic iron chelators and rabbit reticulocytes. Biochim Biophys Acta *373*, 84-99.

Hemmaplardh, D., and Morgan, E.H. (1974b). Transferrin and iron uptake by human cells in culture. Exp Cell Res *87*, 207-212.

Hou, S., Larsen, R.W., Boudko, D., Riley, C.W., Karatan, E., Zimmer, M., Ordal, G.W., and Alam, M. (2000). Myoglobin-like aerotaxis transducers in Archaea and Bacteria. Nature *403*, 540-544.

Hoy, J.A., Robinson, H., Trent, J.T., 3rd, Kakar, S., Smagghe, B.J., and Hargrove, M.S. (2007). Plant hemoglobins: a molecular fossil record for the evolution of oxygen transport. J Mol Biol *371*, 168-179.

Hrkal, Z., and Vodrazka, Z. (1967). A study of the conformation of human globin in solution by optical methods. Biochim Biophys Acta *133*, 527-534.

Hughson, F.M., and Baldwin, R.L. (1989). Use of site-directed mutagenesis to destabilize native apomyoglobin relative to folding intermediates. Biochemistry *28*, 4415-4422.

Hughson, F.M., Barrick, D., and Baldwin, R.L. (1991). Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. Biochemistry *30*, 4113-4118.

Hughson, F.M., Wright, P.E., and Baldwin, R.L. (1990). Structural characterization of a partly folded apomyoglobin intermediate. Science *249*, 1544-1548.

Hunefeld, F.L. (1840). Die Chemismus in der tierischer Organization (Leipzig).

Iannuzzi, C., Vilasi, S., Portaccio, M., Irace, G., and Sirangelo, I. (2007). Heme binding inhibits the fibrillization of amyloidogenic apomyoglobin and determines lack of aggregate cytotoxicity. Protein Sci *16*, 507-516.

Igarashi, K., and Sun, J. (2006). The heme-Bach1 pathway in the regulation of oxidative stress response and erythroid differentiation. Antioxid Redox Signal *8*, 107-118.

Irace, G., Balestrieri, C., Parlato, G., Servillo, L., and Colonna, G. (1981). Tryptophanyl fluorescence heterogeneity of apomyoglobins. Correlation with the presence of two distinct structural domains. Biochemistry *20*, 792-799.

Jamin, M. (2005). The folding process of apomyoglobin. Protein Pept Lett *12*, 229-234.

Jamin, M., and Baldwin, R.L. (1998). Two forms of the pH 4 folding intermediate of apomyoglobin. J Mol Biol *276*, 491-504.

Jeney, V., Balla, J., Yachie, A., Varga, Z., Vercellotti, G.M., Eaton, J.W., and Balla, G. (2002). Pro-oxidant and cytotoxic effects of circulating heme. Blood *100*, 879-887.

Jennings, P.A., and Wright, P.E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. Science *262*, 892-896.

Kakar, S., Hoffman, F.G., Storz, J.F., Fabian, M., and Hargrove, M.S. (2010). Structure and reactivity of hexacoordinate hemoglobins. Biophys Chem.

Kan, Y.W. (1985). Molecular pathology of alpha-thalassemia. Ann N Y Acad Sci *445*, 28-36.

Kawamura-Konishi, Y., Kihara, H., and Suzuki, H. (1988). Reconstitution of myoglobin from apoprotein and heme, monitored by stopped-flow absorption, fluorescence and circular dichroism. Eur J Biochem *170*, 589-595.

Kelly, J.W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. Curr Opin Struct Biol *8*, 101-106.

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature *181*, 662-666.

Kihm, A.J., Kong, Y., Hong, W., Russell, J.E., Rouda, S., Adachi, K., Simon, M.C., Blobel, G.A., and Weiss, M.J. (2002). An abundant erythroid protein that stabilizes free alpha-haemoglobin. Nature *417*, 758-763.

Kirby, E.P., and Steiner, R.F. (1970). The tryptophan microenvironments in apomyoglobin. J Biol Chem *245*, 6300-6306.

Knochel, J.P. (1993). Mechanisms of rhabdomyolysis. Curr Opin Rheumatol *5*, 725-731.

Komar, A.A., Kommer, A., Krasheninnikov, I.A., and Spirin, A.S. (1993). Cotranslational heme binding to nascent globin chains. FEBS Lett *326*, 261-263.

Kooyman, G.L., and Ponganis, P.J. (1998). The physiological basis of diving to depth: birds and mammals. Annu Rev Physiol *60*, 19-32.

Kundu, S., Trent, J.T., 3rd, and Hargrove, M.S. (2003). Plants, humans and hemoglobins. Trends Plant Sci *8*, 387-393.

Kuzelova, K., Mrhalova, M., and Hrkal, Z. (1997). Kinetics of heme interaction with heme-binding proteins: the effect of heme aggregation state. Biochim Biophys Acta *1336*, 497-501.

Landfried, D.A., Vuletich, D.A., Pond, M.P., and Lecomte, J.T. (2007). Structural and thermodynamic consequences of b heme binding for monomeric apoglobins and other apoproteins. Gene *398*, 12-28.

Leckner, J., Bonander, N., Wittung-Stafshede, P., Malmstrom, B.G., and Karlsson, B.G. (1997). The effect of the metal ion on the folding energetics of azurin: a comparison of the native, zinc and apoprotein. Biochim Biophys Acta *1342*, 19-27.

Lecomte, J.T., Sukits, S.F., Bhattacharya, S., and Falzone, C.J. (1999). Conformational properties of native sperm whale apomyoglobin in solution. Protein Sci *8*, 1484-1491.

Lee, V.W., Chen, Y.L., and Konermann, L. (1999). Reconstitution of acid-denatured holomyoglobin studied by time-resolved electrospray ionization mass spectrometry. Anal Chem *71*, 4154-4159.

Leutzinger, Y., and Beychok, S. (1981). Kinetics and mechanism of heme-induced refolding of human alpha-globin. Proc Natl Acad Sci U S A *78*, 780-784.

Levere, R.D., and Granick, S. (1965). Control of hemoglobin synthesis in the cultured chick blastoderm by delta-aminolevulinic acid synthetase: increase in the rate of hemoglobin formation with delta-aminolevulinic acid. Proc Natl Acad Sci U S A *54*, 134-137.

Light, W.R. (1987). Interactions of Heme with Apomyoglobin and Lipid Bilayers. Ph.D. Thesis, Rice University, Houston, TX.

Light, W.R., Rohlfs, R.J., Palmer, G., and Olson, J.S. (1987). Functional effects of heme orientational disorder in sperm whale myoglobin. J Biol Chem *262*, 46-52.

Liong, E.C., Dou, Y., Scott, E.E., Olson, J.S., and Phillips, G.N., Jr. (2001). Waterproofing the heme pocket. Role of proximal amino acid side chains in preventing hemin loss from myoglobin. J Biol Chem *276*, 9093-9100.

Loh, S.N., Kay, M.S., and Baldwin, R.L. (1995). Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. Proc Natl Acad Sci U S A *92*, 5446-5450.

Marden, M.C., Hazard, E.S., Leclerc, L., and Gibson, Q.H. (1989). Flash photolysis of the serum albumin-heme-CO complex. Biochemistry *28*, 4422-4426.

Marks, J., Pozdnyakova, I., Guidry, J., and Wittung-Stafshede, P. (2004). Methionine-121 coordination determines metal specificity in unfolded Pseudomonas aeruginosa azurin. J Biol Inorg Chem *9*, 281-288.

Merx, M.W., Godecke, A., Flogel, U., and Schrader, J. (2005). Oxygen supply and nitric oxide scavenging by myoglobin contribute to exercise endurance and cardiac function. Faseb J *19*, 1015-1017.

Messerschmidt, A. (2001). Handbook of metalloproteins (New York, Wiley).

Miranda, J.J., Maillett, D.H., Soman, J., and Olson, J.S. (2005). Thermoglobin, oxygen-avid hemoglobin in a bacterial hyperthermophile. J Biol Chem *280*, 36754-36761.

Mitchell, D.T., Ernst, S.R., Wu, W.X., and Hackert, M.L. (1995). Three-dimensional structure of a hemichrome hemoglobin from Caudina arenicola. Acta Crystallogr D Biol Crystallogr *51*, 647-653.

Moczygemba, C., Guidry, J., and Wittung-Stafshede, P. (2000). Heme orientation affects holo-myoglobin folding and unfolding kinetics. FEBS Lett *470*, 203-206.

Mollan, T.L., Yu, X., Weiss, M.J., and Olson, J.S. (2010). The role of alpha-hemoglobin stabilizing protein in redox chemistry, denaturation, and hemoglobin assembly. Antioxid Redox Signal *12(2)*, 219-231.

Naka, T., Jones, D., Baldwin, I., Fealy, N., Bates, S., Goehl, H., Morgera, S., Neumayer, H.H., and Bellomo, R. (2005). Myoglobin clearance by super high-flux hemofiltration in a case of severe rhabdomyolysis: a case report. Crit Care *9*, R90-95.

Needham, M., and Mastaglia, F.L. (2007). Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches. Lancet Neurol *6*, 620-631.

Nicolas, G., Bennoun, M., Porteu, A., Mativet, S., Beaumont, C., Grandchamp, B., Sirito, M., Sawadogo, M., Kahn, A., and Vaulont, S. (2002). Severe iron deficiency anemia in transgenic mice expressing liver hepcidin. Proc Natl Acad Sci U S A *99*, 4596-4601.

Nishii, I., Kataoka, M., and Goto, Y. (1995). Thermodynamic stability of the molten globule states of apomyoglobin. J Mol Biol *250*, 223-238.

Nishii, I., Kataoka, M., Tokunaga, F., and Goto, Y. (1994). Cold denaturation of the molten globule states of apomyoglobin and a profile for protein folding. Biochemistry *33*, 4903-4909.

Nishimura, C., Dyson, H.J., and Wright, P.E. (2006). Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin. J Mol Biol *355*, 139-156.

Nishimura, C., Wright, P.E., and Dyson, H.J. (2003). Role of the B helix in early folding events in apomyoglobin: evidence from site-directed mutagenesis for native-like long range interactions. J Mol Biol *334*, 293-307.

Olson, J.S. (2005). Design of heme protein-based blood substitutes (NIH-grant application. Houston, TX).

Olson, J.S., and Phillips, G.N., Jr. (1997). Myoglobin discriminates between O_2, NO, and CO by electrostatic interactions with the bound ligand. J Biol Inorg Chem *2*, 544-542.

Ordway, G.A., and Garry, D.J. (2004). Myoglobin: an essential hemoprotein in striated muscle. J of Exp Bio *207*, 3441-3446.

Orkin, S.H. (1995). Hematopoiesis: how does it happen? Curr Opin Cell Biol *7*, 870-877.

Pang, L., Weiss, M.J., and Poncz, M. (2005). Megakaryocyte biology and related disorders. J Clin Invest *115*, 3332-3338.

Pesce, A., Nardini, M., Ascenzi, P., Geuens, E., Dewilde, S., Moens, L., Bolognesi, M., Riggs, A.F., Hale, A., Deng, P., *et al.* (2004). Thr-E11 regulates O2 affinity in Cerebratulus lacteus mini-hemoglobin. J Biol Chem *279*, 33662-33672.

Pesce, A., Nardini, M., Dewilde, S., Ascenzi, P., Riggs, A.F., Yamauchi, K., Geuens, E., Moens, L., and Bolognesi, M. (2001). Crystallization and preliminary X-ray analysis of neural haemoglobin from the nemertean worm Cerebratulus lacteus. Acta Crystallogr D Biol Crystallogr *57*, 1897-1899.

Pesce, A., Nardini, M., Dewilde, S., Geuens, E., Yamauchi, K., Ascenzi, P., Riggs, A.F., Moens, L., and Bolognesi, M. (2002). The 109 residue nerve tissue minihemoglobin from Cerebratulus lacteus highlights striking structural plasticity of the alpha-helical globin fold. Structure *10*, 725-735.

Piccard, H., Van den Steen, P.E., and Opdenakker, G. (2007). Hemopexin domains as multifunctional liganding modules in matrix metalloproteinases and other proteins. J Leukoc Biol *81*, 870-892.

Ponka, P. (1997). Tissue-specific regulation of iron metabolism and heme synthesis: distinct control mechanisms in erythroid cells. Blood *89*, 1-25.

Ponka, P. (1999). Cell biology of heme. Am J Med Sci *318*, 241-256.

Quillin, M.L., Arduini, R.M., Olson, J.S., and Phillips, G.N., Jr. (1993). High-resolution crystal structures of distal histidine mutants of sperm whale myoglobin. J Mol Biol *234*, 140-155.

Rachmilewitz, E.A., Peisach, J., and Blumberg, W.E. (1971). Studies on the stability of oxyhemoglobin A and its constituent chains and their derivatives. J Biol Chem *246*, 3356-3366.

Raju, V.S., and Maines, M.D. (1996). Renal ischemia/reperfusion up-regulates heme oxygenase-1 (HSP32) expression and increases cGMP in rat heart. J Pharmacol Exp Ther *277*, 1814-1822.

Ramos, C.H., Kay, M.S., and Baldwin, R.L. (1999). Putative interhelix ion pairs involved in the stability of myoglobin. Biochemistry *38*, 9783-9790.

Ramsay, G., Ionescu, R., and Eftink, M.R. (1995). Modified spectrophotometer for multi-dimensional circular dichroism/fluorescence data acquisition in titration experiments: application to the pH and guanidine-HCI induced unfolding of apomyoglobin. Biophys J *69*, 701-707.

Rassaf, T., Flogel, U., Drexhage, C., Hendgen-Cotta, U., Kelm, M., and Schrader, J. (2007). Nitrite reductase function of deoxymyoglobin: oxygen sensor and regulator of cardiac energetics and function. Circ Res *100*, 1749-1754.

Reeder, B.J. (2010). The redox activity of hemoglobins: from physiologic functions to pathologic mechanisms. Antioxid Redox Signal *13*, 1087-1123.

Reeder, B.J., and Wilson, M.T. (2005). Hemoglobin and myoglobin associated oxidative stress: from molecular mechanisms to disease States. Curr Med Chem *12*, 2741-2751.

Regis, W.C., Fattori, J., Santoro, M.M., Jamin, M., and Ramos, C.H. (2005). On the difference in stability between horse and sperm whale myoglobins. Arch Biochem Biophys *436*, 168-177.

Reisberg, P. (1980). Mechanisms of Ligand Binding to Hemoglobin. Ph.D. Thesis Rice University.

Reynafarje, B. (1962a). Myoglobin content and enzymatic activity of human skeletal muscle--their relation with the process of adaptation to high altitude. Tech Doc Rep SAMTDR USAF Sch Aerosp Med *SAM-TDR-62-89*, 8p.

Reynafarje, B. (1962b). Myoglobin content and enzymatic activity of muscle and altitude adaptation. J Appl Physiol *17*, 301-305.

Rifkind, J.M., Abugo, O., Levy, A., and Heim, J. (1994). Detection, formation, and relevance of hemichromes and hemochromes. Methods Enzymol *231*, 449-480.

Robach, P., Cairo, G., Gelfi, C., Bernuzzi, F., Pilegaard, H., Vigano, A., Santambrogio, P., Cerretelli, P., Calbet, J.A., Moutereau, S., *et al.* (2007). Strong iron demand during hypoxia-induced erythropoiesis is associated with down-regulation of iron-related proteins and myoglobin in human skeletal muscle. Blood *109*, 4724-4731.

Robinson, C.R., Liu, Y., O'Brien, R., Sligar, S.G., and Sturtevant, J.M. (1998). A differential scanning calorimetric study of the thermal unfolding of apo- and holo-cytochrome b562. Protein Sci *7*, 961-965.

Robinson, C.R., Liu, Y., Thomson, J.A., Sturtevant, J.M., and Sligar, S.G. (1997). Energetics of heme binding to native and denatured states of cytochrome b562. Biochemistry *36*, 16141-16146.

Romero-Herrera, A.E., and Lehmann, H. (1974). The amino acid sequence of human myoglobin and its minor fractions. Proc R Soc Lond B Biol Sci *186*, 249-279.

Rose, M.Y., and Olson, J.S. (1983). The kinetic mechanism of heme binding to human apohemoglobin. J Biol Chem *258*, 4298-4303.

Rossi-Fanelli, A., Antonini, E., and Caputo, A. (1964). Hemoglobin and Myoglobin. Advan Protein Chem *19*, 73-222.

Rouault, T.A. (2004). Microbiology. Pathogenic bacteria prefer heme. Science *305*, 1577-1578.

Royer, W.E., Jr., Strand, K., van Heel, M., and Hendrickson, W.A. (2000). Structural hierarchy in erythrocruorin, the giant respiratory assemblage of annelids. Proc Natl Acad Sci U S A *97*, 7107-7111.

Sackmann, E. (1995). Biological membranes Architecture and Function

Handbook of Biological Physics, Vol 1 (Elsevier).

Salter, M.D., Nienhaus, K., Nienhaus, G.U., Dewilde, S., Moens, L., Pesce, A., Nardini, M., Bolognesi, M., and Olson, J.S. (2008). The apolar channel in *Cerebratulus lacteus* hemoglobin is the route for O_2 entry and exit. J Biol Chem *283*, 35689-35702.

Sanchez-Ruiz, J.M., Lopez-Lacomba, J.L., Cortijo, M., and Mateo, P.L. (1988). Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. Biochemistry *27*, 1648-1652.

Scarabelli, T.M., and Gottlieb, R.A. (2004). Functional and clinical repercussions of myocyte apoptosis in the multifaceted damage by ischemia/reperfusion injury: old and new concepts after 10 years of contributions. Cell Death Differ *11 Suppl 2*, S144-152.

Schechter, A.N. (2008). Hemoglobin research and the origins of molecular medicine. Blood *112*, 3927-3938.

Schlieper, G., Kim, J.H., Molojavyi, A., Jacoby, C., Laussmann, T., Flogel, U., Godecke, A., and Schrader, J. (2004). Adaptation of the myoglobin knockout mouse to hypoxic stress. Am J Physiol Regul Integr Comp Physiol *286*, R786-792.

Schmidt, M., Gerlach, F., Avivi, A., Laufs, T., Wystub, S., Simpson, J.C., Nevo, E., Saaler-Reinhardt, S., Reuss, S., Hankeln, T., *et al.* (2004). Cytoglobin is a respiratory protein in connective tissue and neurons, which is up-regulated by hypoxia. J Biol Chem *279*, 8063-8069.

Schreiter, E.R., Rodriguez, M.M., Weichsel, A., Montfort, W.R., and Bonaventura, J. (2007). S-nitrosylation-induced conformational change in blackfin tuna myoglobin. J Biol Chem *282*, 19773-19780.

Scott, E.E., Paster, E.V., and Olson, J.S. (2000). The stabilities of mammalian apomyoglobins vary over a 600-fold range and can be enhanced by comparative mutagenesis. J Biol Chem *275*, 27129-27136.

Scott, M.D., van den Berg, J.J., Repka, T., Rouyer-Fessard, P., Hebbel, R.P., Beuzard, Y., and Lubin, B.H. (1993). Effect of excess alpha-hemoglobin chains on cellular and membrane oxidation in model beta-thalassemic erythrocytes. J Clin Invest *91*, 1706-1712.

Scouloudi, H., and Baker, E.N. (1978). X-ray crystallographic studies of seal myoglobin. The molecule at 2.5 A resolution. J Mol Biol *126*, 637-660.

Severinghaus, J.W., and Astrup, P.B. (1986). History of blood gas analysis. VI. Oximetry. J Clin Monit *2*, 270-288.

Shack, J., and Clark, W.M. (1947). Metalloporphyrins. J Biol Chem *171*, 143-187.

Shaeffer, J.R. (1967). Evidence for soluble alpha-chains as intermediates in hemoglobin synthesis in the rabbit reticulocyte. Biochem Biophys Res Commun *28*, 647-652.

Sigurdsson, E.M., Wisniewski, T., and Frangione, B. (2002). Infectivity of amyloid diseases. Trends Mol Med *8*, 411-413.

Singh, S., Manda, S.M., Sikder, D., Birrer, M.J., Rothermel, B.A., Garry, D.J., and Mammen, P.P. (2009). Calcineurin activates cytoglobin transcription in hypoxic myocytes. J Biol Chem *284*, 10409-10421.

Slater, M.S., and Mullins, R.J. (1998). Rhabdomyolysis and myoglobinuric renal failure in trauma and surgical patients: a review. J Am Coll Surg *186*, 693-716.

Smerdon, S.J., Oldfield, T.J., Dodson, E.J., Dodson, G.G., Hubbard, R.E., and Wilkinson, A.J. (1990). Determination of the crystal structure of recombinant pig myoglobin by molecular replacement and its refinement. Acta Crystallogr B *46 (Pt 3)*, 370-377.

Smith, L. (2003). The effects of amino acid substitution on apomyoglobin stability, folding intermediates, and holoprotein expression. Ph.D. Thesis, Rice University, Houston, TX.

Socolovsky, M. (2007). Molecular insights into stress erythropoiesis. Curr Opin Hematol *14*, 215-224.

Springer, B.A., Egeberg, K.D., Sligar, S.G., Rohlfs, R.J., Mathews, A.J., and Olson, J.S. (1989). Discrimination between oxygen and carbon monoxide and inhibition of autooxidation by myoglobin. Site-directed mutagenesis of the distal histidine. J Biol Chem *264*, 3057-3060.

Springer, B.A., and Sligar, S.G. (1987). High-level expression of sperm whale myoglobin in Escherichia coli. Proc Natl Acad Sci U S A *84*, 8961-8965.

Steensma, D.P., Gibbons, R.J., and Higgs, D.R. (2005). Acquired alpha-thalassemia in association with myelodysplastic syndrome and other hematologic malignancies. Blood *105*, 443-452.

Stojiljkovic, I., and Perkins-Balding, D. (2002). Processing of heme and heme-containing proteins by bacteria. DNA Cell Biol *21*, 281-295.

Strand, K., Knapp, J.E., Bhyravbhatla, B., and Royer, W.E., Jr. (2004). Crystal structure of the hemoglobin dodecamer from Lumbricus erythrocruorin: allosteric core of giant annelid respiratory complexes. J Mol Biol *344*, 119-134.

Taketani, S. (2005). Aquisition, mobilization and utilization of cellular iron and heme: endless findings and growing evidence of tight regulation. Tohoku J Exp Med *205*, 297-318.

Tavill, A.S., Grayzel, A.I., Vanderhoff, G.A., and London, I.M. (1967). The control of hemoglobin synthesis. Trans Assoc Am Physicians *80*, 305-313.

Tcherkasskaya, O., Bychkova, V.E., Uversky, V.N., and Gronenborn, A.M. (2000a). Multisite fluorescence in proteins with multiple tryptophan residues. Apomyoglobin natural variants and site-directed mutants. J Biol Chem *275*, 36285-36294.

Tcherkasskaya, O., Ptitsyn, O.B., and Knutson, J.R. (2000b). Nanosecond dynamics of tryptophans in different conformational states of apomyoglobin proteins. Biochemistry *39*, 1879-1889.

Trevaskis, B., Watts, R.A., Andersson, C.R., Llewellyn, D.J., Hargrove, M.S., Olson, J.S., Dennis, E.S., and Peacock, W.J. (1997). Two hemoglobin genes in Arabidopsis thaliana: the evolutionary origins of leghemoglobins. Proc Natl Acad Sci U S A *94*, 12230-12234.

Ullmann, B.D., Myers, H., Chiranand, W., Lazzell, A.L., Zhao, Q., Vega, L.A., Lopez-Ribot, J.L., Gardner, P.R., and Gustin, M.C. (2004). Inducible defense mechanism against nitric oxide in Candida albicans. Eukaryot Cell *3*, 715-723.

Uzawa, T., Akiyama, S., Kimura, T., Takahashi, S., Ishimori, K., Morishima, I., and Fujisawa, T. (2004). Collapse and search dynamics of apomyoglobin folding revealed by submillisecond observations of alpha-helical content and compactness. Proc Natl Acad Sci U S A *101*, 1171-1176.

Vandergon, T.L., Riggs, C.K., Gorr, T.A., Colacino, J.M., and Riggs, A.F. (1998). The mini-hemoglobins in neural and body wall tissue of the nemertean worm, Cerebratulus lacteus. J Biol Chem *273*, 16998-17011.

Vergara, A., Franzese, M., Merlino, A., Bonomi, G., Verde, C., Giordano, D., di Prisco, G., Lee, H.C., Peisach, J., and Mazzarella, L. (2009). Correlation between

hemichrome stability and the root effect in tetrameric hemoglobins. Biophys J *97*, 866-874.

Vergara, A., Franzese, M., Merlino, A., Vitagliano, L., Verde, C., di Prisco, G., Lee, H.C., Peisach, J., and Mazzarella, L. (2007). Structural characterization of ferric hemoglobins from three antarctic fish species of the suborder notothenioidei. Biophys J *93*, 2822-2829.

Vergara, A., Vitagliano, L., Verde, C., di Prisco, G., and Mazzarella, L. (2008). Spectroscopic and crystallographic characterization of bis-histidyl adducts in tetrameric hemoglobins. Methods Enzymol *436*, 425-444.

Vinogradov, S.N., Hoogewijs, D., Bailly, X., Mizuguchi, K., Dewilde, S., Moens, L., and Vanfleteren, J.R. (2007). A model of globin evolution. Gene *398*, 132-142.

Vinogradov, S.N., Shlom, J.M., Hall, B.C., Kapp, O.H., and Mizukami, H. (1977). The dissociation of Lumbricus terrestris hemoglobin: a model of its subunit structure. Biochim Biophys Acta *492*, 136-155.

Wagener, F.A., Eggert, A., Boerman, O.C., Oyen, W.J., Verhofstad, A., Abraham, N.G., Adema, G., van Kooyk, Y., de Witte, T., and Figdor, C.G. (2001). Heme is a potent inducer of inflammation in mice and is counteracted by heme oxygenase. Blood *98*, 1802-1811.

Waks, M., Yip, Y.K., and Beychok, S. (1973). Influence of prosthetic groups on protein folding and subunit assembly. Recombination of separated human alpha-and beta-globin chains with heme and alloplex interactions of globin chains with heme-containing subunits. J Biol Chem *248*, 6462-6470.

Weiss, M.J., Zhou, S., Feng, L., Gell, D.A., Mackay, J.P., Shi, Y., and Gow, A.J. (2005). Role of alpha-hemoglobin-stabilizing protein in normal erythropoiesis and beta-thalassemia. Ann N Y Acad Sci *1054*, 103-117.

Westermark, P., Benson, M.D., Buxbaum, J.N., Cohen, A.S., Frangione, B., Ikeda, S., Masters, C.L., Merlini, G., Saraiva, M.J., and Sipe, J.D. (2002). Amyloid fibril protein nomenclature -- 2002. Amyloid *9*, 197-200.

Wilson, C.J., Apiyo, D., and Wittung-Stafshede, P. (2004). Role of cofactors in metalloprotein folding. Q Rev Biophys *37*, 285-314.

Winterbourn, C.C. (1990). Oxidative denaturation in congenital hemolytic anemias: the unstable hemoglobins. Semin Hematol *27*, 41-50.

Winterbourn, C.C., and Carrell, R.W. (1974). Studies of hemoglobin denaturation and Heinz body formation in the unstable hemoglobins. J Clin Invest *54*, 678-689.

Winterhalter, K.H., Amiconi, G., and Antonini, E. (1968). Functional properties of a hemoglobin carrying heme only on alpha chains. Biochemistry *7*, 2228-2232.

Winterhalter, K.H., and Colosimo, A. (1971). Chromatographic isolation and characterization of isolated chains from hemoglobin after regeneration of sulfhydryl groups. Biochemistry *10*, 621-624.

Winterhalter, K.H., and Deranleau, D.A. (1967). The structure of a hemoglobin carrying only two hemes. Biochemistry *6*, 3136-3143.

Winterhalter, K.H., and Huehns, E.R. (1964). Preparations, Properties, and Specific Recombination of Alpha-Beta-Globin Subunits. J Biol Chem *239*, 3699-3705.

Wittenberg, J.B. (1970). Myoglobin-facilitated oxygen diffusion: role of myoglobin in oxygen entry into muscle. Physiol Rev *50*, 559-636.

Wittenberg, J.B., and Wittenberg, B.A. (2003). Myoglobin function reassessed. J Exp Biol *206*, 2011-2020.

Wittung-Stafshede, P., Malmstrom, B.G., Winkler, J.R., and Gray, H.B. (1998). Folding of deoxymyoglobin triggered by electron transfer. Journal of Physical Chemistry *102*, 5599-5601.

Xu, F., Quandt, K.S., and Hultquist, D.E. (1992). Characterization of NADPH-dependent methemoglobin reductase as a heme-binding protein present in erythrocytes and liver. Proc Natl Acad Sci U S A *89*, 2130-2134.

Yip, Y.K., Waks, M., and Beychok, S. (1972). Influence of prosthetic groups on protein folding and subunit assembly. I. Conformational differences between separated human alpha- and beta- globins. J Biol Chem *247*, 7237-7244.

Yu, X., Kong, Y., Dore, L.C., Abdulmalik, O., Katein, A.M., Zhou, S., Choi, J.K., Gell, D., Mackay, J.P., Gow, A.J., *et al.* (2007). An erythroid chaperone that facilitates folding of alpha-globin subunits for hemoglobin synthesis. J Clin Invest *117*, 1856-1865.

Zager, R.A. (1996). Rhabdomyolysis and myohemoglobinuric acute renal failure. Kidney Int *49*, 314-326.

Zhang, W., Olson, J.S., and Phillips, G.N., Jr. (2005). Biophysical and kinetic characterization of HemAT, an aerotaxis receptor from Bacillus subtilis. Biophys J *88*, 2801-2814.

Zhu, J., and Emerson, S.G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. Oncogene *21*, 3295-3313.

Zwanzig, R., Szabo, A., and Bagchi, B. (1992). Levinthal's paradox. Proc Natl Acad Sci U S A *89*, 20-22.

APPENDIX

Derivations of equations and computer programs

A.1. Simple mechanism for monomeric globin unfolding

$$K_{NH} = \frac{[N][H]}{[NH]} = K_{NH}^{0} \exp\left(\frac{m_{NH}[GuHCl]}{RT}\right) \tag{A.1.1}$$

$$K_{NU} = \frac{[U]}{[N]} = K_{NU}^{0} \exp\left(\frac{m_{NU}[GuHCl]}{RT}\right) \tag{A.1.2}$$

$$P_0 = [NH] + [N] + [U] = [NH]\left(1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \tag{A.1.3}$$

$$[NH] = P_0 Y_{NH} = P_0\left(\frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}}\right) \tag{A.1.4}$$

$$[H] = [N] + [U] = [mHb]\left(\frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) = P_0\left(\frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}}\right)\left(\frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \tag{A.1.5}$$

$$[H] = \frac{-\left(K_{NH} + K_{NH}K_{NU}\right) + \sqrt{\left(K_{NH} + K_{NH}K_{NU}\right)^2 + 4P_0\left(K_{NH} + K_{NH}K_{NU}\right)}}{2} \tag{A.1.6}$$

The population fractions of the three species can be obtained with the following equations:

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}} \tag{A.1.7}$$

$$Y_N = \frac{\frac{K_{NH}}{[H]}}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}} \tag{A.1.8}$$

$$Y_U = \frac{\frac{K_{NH}K_{NU}}{[H]}}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}} \tag{A.1.9}$$

A.2. Simple mechanism for monomeric globin unfolding with hemin binding to the unfolded state

$$K_{NH} = \frac{[N][H]}{[mHb]} = K_{NH}^{0} \exp\left(\frac{m_{NH}[GuHCl]}{RT}\right) \tag{A.2.1}$$

$$K_{UH} = \frac{[U][H]}{[UH]} = K_{UH}^{0} \exp\left(\frac{m_{UH}[GuHCl]}{RT}\right) \tag{A.2.2}$$

$$K_{NU} = \frac{[U]}{[N]} = K_{NU}^{0} \exp\left(\frac{m_{NU}[GuHCl]}{RT}\right) \tag{A.2.3}$$

$$P_0 = [NH] + [N] + [U] + [UH] = [NH]\left(1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]} + \frac{K_{NH}K_{NU}}{K_{UH}}\right) \tag{A.2.4}$$

$$\begin{aligned} [H] &= [N] + [U] = [mHb]\left(\frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \\ &= P_0\left(\frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]} + \frac{K_{NH}K_{NU}}{K_{UH}}}\right)\left(\frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \end{aligned} \tag{A.2.5}$$

The concentration of hemin [H] is obtained using the following expression:

$$[H] = \frac{-\left(K_{NH} + K_{NH}K_{NU}\right) + \sqrt{\left(K_{NH} + K_{NH}K_{NU}\right)^2 + 4P_0\left(1 + \frac{K_{NH}K_{NU}}{K_{UH}}\right)\left(K_{NH} + K_{NH}K_{NU}\right)}}{2\left(1 + \frac{K_{NH}K_{NU}}{K_{UH}}\right)} \tag{A.2.6}$$

The population fractions of the three species can be obtained with the following equations:

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]} + \frac{K_{NH}K_{NU}}{K_{UH}}} \tag{A.2.7}$$

A.3. Simple mechanism for monomeric globin unfolding with hemin dimerization

$$K_{NH} = \frac{[N][H]}{[mHb]} = K_{NH}^{0} \exp\left(\frac{m_{NH}[GuHCl]}{RT}\right) \tag{A.3.1}$$

$$K_{NU} = \frac{[U]}{[N]} = K_{NU}^{0} \exp\left(\frac{m_{NU}[GuHCl]}{RT}\right) \tag{A.3.2}$$

$$K_{H_2} = \frac{[H_2]}{[H]^2} = K_{H_2}^{0} \exp\left(\frac{m_{H_2}[GuHCl]}{RT}\right) \tag{A.3.3}$$

$$P_0 = [NH] + [N] + [U] = [NH]\left(1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \tag{A.3.4}$$

$$[N] + [U] = [NH]\left(\frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) \tag{A.3.5}$$

$$[H] + 2[H_2] = [H] + 2K_{H_2}[H]^2 \tag{A.3.6}$$

$$P_0\left(\frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}}\right)\left(\frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}\right) = [H] + 2K_{H_2}[H]^2 \tag{A.3.7}$$

Equation A.3.7 can be rearranged into a cubic equation with the coefficients being:

$$a = 2K_{H_2} \tag{A.3.8}$$

$$b = 1 + 2K_{H_2}\left(K_{NH} + K_{NH}K_{NU}\right) \tag{A.3.9}$$

$$c = K_{NH} + K_{NH}K_{NU} \tag{A.3.10}$$

$$d = -P_0\left(K_{NH} + K_{NH}K_{NU}\right) \tag{A.3.11}$$

The appropriate root for [H] can be solved using Cardano's method:

$$[H] = -\frac{b}{3a}$$

$$-\frac{1}{3a}\left(\sqrt[3]{\frac{2b^2 - 9abc + 27a^2d + \sqrt{\left(2b^2 - 9abc + 27a^2d\right)^2 - 4\left(b^2 - 3ac\right)^2}}{2}}\right)$$

$$-\frac{1}{3a}\left(\sqrt[3]{\frac{2b^2 - 9abc + 27a^2d - \sqrt{\left(2b^2 - 9abc + 27a^2d\right)^2 - 4\left(b^2 - 3ac\right)^2}}{2}}\right) \qquad \text{(A.3.12)}$$

Then the population fractions can be computed as in Appendix A.1:

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NU}}{[H]}} \qquad \text{(A.3.13)}$$

A.4. Three-state mechanism for the unfolding of apoMb

The N-to-I and I-to-U transitions are defined according to:

$$K_{NI} = \frac{[I]}{[N]} = K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) \tag{A.4.1}$$

$$K_{IU} = \frac{[U]}{[I]} = K_{IU}^{0} \exp\left(\frac{m_{IU}[GuHCl]}{RT}\right) \tag{A.4.2}$$

The concentrations of the I and U states are then defined as:

$$[I] = [N]K_{NI} \tag{A.4.3}$$

$$[U] = [N]K_{NI}K_{IU} \tag{A.4.4}$$

The observable can be computed as:

$$S = \frac{S_N + S_I K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) + S_U K_{NI}^{0} K_{IU}^{0} \exp\left(\frac{(m_{NI} + m_{IU})[GuHCl]}{RT}\right)}{1 + K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) + K_{NI}^{0} K_{IU}^{0} \exp\left(\frac{(m_{NI} + m_{IU})[GuHCl]}{RT}\right)} \tag{A.4.5}$$

A.5. Six-state unfolding mechanism or holoMb

The dependence of [GuHCl] on the stability and dissociation constants is expressed as:

$$K_{NI} = \frac{[I]}{[N]} = K_{NI}^{0} \exp\left(\frac{m_{NI}[GuHCl]}{RT}\right) \tag{A.5.1}$$

$$K_{IU} = \frac{[U]}{[I]} = K_{IU}^{0} \exp\left(\frac{m_{IU}[GuHCl]}{RT}\right) \tag{A.5.2}$$

$$K_{NH} = \frac{[N][H]}{[NH]} = K_{NH}^{0} \exp\left(\frac{m_{NH}[GuHCl]}{RT}\right) \tag{A.5.3}$$

$$K_{IH} = \frac{[I][H]}{[IH]} = K_{IH}^{0} \exp\left(\frac{m_{IH}[GuHCl]}{RT}\right) \tag{A.5.4}$$

$$K_{UH} = \frac{[U][H]}{[UH]} = K_{UH}^{0} \exp\left(\frac{m_{UH}[GuHCl]}{RT}\right) \tag{A.5.6}$$

$$[H] = [N] + [I] + [U] \tag{A.5.7}$$

$$\frac{[NH]}{[H]}\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right) = [H] \tag{A.5.8}$$

$$[NH] = P_0 Y_{NH} \tag{A.5.9}$$

$$\frac{P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)}{\left(1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NI}}{[H]} + \frac{K_{NH}K_{NI}K_{IU}}{[H]} + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)} = [H]^2 \tag{A.5.10}$$

Equation A.5.10 can be rearranged into a quadratic equation:

$$a[H]^2 + b[H] + c \tag{A.5.11}$$

where:

$$a = \left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right) \tag{A.5.12}$$

$$b = \left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right) \tag{A.5.13}$$

$$c = -P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right) \tag{A.5.14}$$

The only possible solution for [H] is:

$$[H] = \frac{-\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)}{2\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)} + \frac{\sqrt{\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)^2 + 4P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)}}{2\left(1 + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)} \tag{A.5.15}$$

The population fractions can be computed as in Appendix A.

$$Y_{NH} = \frac{1}{1 + \frac{K_{NH}}{[H]} + \frac{K_{NH}K_{NI}}{[H]} + \frac{K_{NH}K_{NI}K_{IU}}{[H]} + \frac{K_{NH}K_{NI}}{K_{IH}} + \frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}} \tag{A.5.16}$$

A.6. Six-state unfolding mechanism for holoMb with hemin dimerization

Using the same stability constants and hemin affinities in the six-state model, the hemin dimerization step is defined as:

$$K_{H_2} = \frac{[H_2]}{[H]^2} = K^0_{H_2} \exp\left(\frac{m_{H_2}[GuHCl]}{RT}\right) \tag{A.6.1}$$

$$[N]+[I]+[U] = [H]+2[D] \tag{A.6.2}$$

$$\frac{[NH]}{[H]}\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right) = [H]+2[H]^2 K_{H_2} \tag{A.6.3}$$

$$\frac{P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right)}{\left(1+\frac{K_{NH}}{[H]}+\frac{K_{NH}K_{NI}}{[H]}+\frac{K_{NH}K_{NI}K_{IU}}{[H]}+\frac{K_{NH}K_{NI}}{K_{IH}}+\frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right)} = [H]^2 + 2K_{H_2}[H]^3 \tag{A.6.4}$$

Equation A.6.4 can be rearranged into a cubic equation:

$$a[H]^3 + b[H]^2 + c[H] + d \tag{A.6.5}$$

where:

$$a = 2K_{H_2}\left(1+\frac{K_{NH}K_{NI}}{K_{IH}}+\frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right) \tag{A.6.6}$$

$$b = 2K_{H_2}(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}) + \left(1+\frac{K_{NH}K_{NI}}{K_{IH}}+\frac{K_{NH}K_{NI}K_{IU}}{K_{UH}}\right) \tag{A.6.7}$$

$$c = K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU} \tag{A.6.8}$$

$$d = -P_0\left(K_{NH} + K_{NH}K_{NI} + K_{NH}K_{NI}K_{IU}\right) \tag{A.6.9}$$

The roots for this complex equation can be obtained using Cardano's method.

A.7. Simple two-state process for the unfolding of the apoHb dimer

The equilibrium folding constant is defined as:

$$K_{2U,D} = \frac{[D]}{[U]^2} = K^0_{2U,D} \exp\left(\frac{-m_{2U,D}[GuHCl]}{RT}\right) \tag{A.7.1}$$

$$[D] = K_{2U,D}[U]^2 \tag{A.7.2}$$

$$P_0 = 2[D] + [U] = 2K_{2U,D}[U]^2 + [U] \tag{A.7.3}$$

The concentration of the unfolded state U is the root of a simple quadratic:

$$[U] = \frac{-1 + \sqrt{1 + 8P_0K_{2U,D}}}{4K_{2U,D}} \tag{A.7.4}$$

The population fraction of the folded dimer is defined as:

$$Y_D = \frac{2[D]}{2[D] + [U]} = \frac{2K_{2U,D}\left(\frac{-1 + \sqrt{1 + 8P_0K_{2U,D}}}{4K_{2U,D}}\right)^2}{2K_{2U,D}\left(\frac{-1 + \sqrt{1 + 8P_0K_{2U,D}}}{4K_{2U,D}}\right)^2 + \left(\frac{-1 + \sqrt{1 + 8P_0K_{2U,D}}}{4K_{2U,D}}\right)} \tag{A.7.5}$$

A.8. Unfolding of the apoHb dimer involving a first dissociation followed by unfolding of folded subunits, via an intermediate

The equilibrium folding constants are defined as:

$$K_{U,I} = \frac{[I]}{[U]} = K^0_{U,I} \exp\left(\frac{-m_{U,I}[GuHCl]}{RT}\right) \quad \text{(A.8.1)}$$

$$K_{I,N} = \frac{[N]}{[I]} = K^0_{I,N} \exp\left(\frac{-m_{I,N}[GuHCl]}{RT}\right) \quad \text{(A.8.2)}$$

$$K_{2N,D} = \frac{[D]}{[N]^2} = K^0_{2N,D} \exp\left(\frac{-m_{2N,D}[GuHCl]}{RT}\right) \quad \text{(A.8.3)}$$

The concentration of each species is:

$$[I] = K_{U,I}[U] \quad \text{(A.8.4)}$$

$$[N] = K_{I,N}K_{U,I}[U] \quad \text{(A.8.5)}$$

$$[D] = K_{2N,D}K_{I,N}{}^2K_{U,I}{}^2[U]^2 \quad \text{(A.8.6)}$$

$$P_0 = 2[D]+[N]+[I]+[U] = 2K_{2N,D}K_{I,N}{}^2K_{U,I}{}^2[U]^2 + K_{I,N}K_{U,I}[U] + K_{U,I}[U] + [U] \quad \text{(A.8.7)}$$

Solving equation A.8.7 for [U] leads to the following root:

$$[U] = \frac{-\left(1 + K_{I,N}K_{U,I} + K_{U,I}\right) + \sqrt{\left(1 + K_{I,N}K_{U,I} + K_{U,I}\right)^2 + 8P_0K_{2N,D}K_{I,N}{}^2K_{U,I}{}^2}}{4K_{2N,D}K_{I,N}{}^2K_{U,I}{}^2} \quad \text{(A.8.8)}$$

The population fractions can then be computed as:

$$Y_D = \frac{2[D]}{2[D]+[N]+[I]+[U]} = \frac{2K_{2N,D}K_{I,N}{}^2K_{U,I}{}^2[U]^2}{2K_{2N,D}K_{I,N}{}^2K_{U,I}{}^2[U]^2 + K_{I,N}K_{U,I}[U] + K_{U,I}[U] + [U]} \quad \text{(A.8.9)}$$

A.9. Unfolding of apoHb dimer involving a dimer intermediate

The equilibrium folding constants are defined as:

$$K_{2U,I_D} = \frac{[I_D]}{[U]^2} = K^0_{2U,I_D} \exp\left(\frac{-m_{2U,I_D}[GuHCl]}{RT}\right) \quad \text{(A.9.1)}$$

$$K_{I_D,D} = \frac{[D]}{[I_D]} = K^0_{I_D,D} \exp\left(\frac{-m_{I_D,D}[GuHCl]}{RT}\right) \quad \text{(A.9.2)}$$

The concentration of each species is:

$$[I_D] = K_{2U,I_D}[U]^2 \quad \text{(A.9.3)}$$

$$[D] = K_{2U,I_D} K_{I_D,D}[U]^2 \quad \text{(A.9.4)}$$

$$P_0 = 2[D] + 2[I_D] + [U] = 2K_{2U,I_D} K_{I_D,D}[U]^2 + 2K_{2U,I_D}[U]^2 + [U] \quad \text{(A.9.5)}$$

Solving equation A.9.7 for [U] leads to the following root:

$$[U] = \frac{-1 + \sqrt{1 + 8P_0\left(K_{2U,I_D} K_{I_D,D} + K_{2U,I_D}\right)}}{4\left(K_{2U,I_D} K_{I_D,D} + K_{2U,I_D}\right)} \quad \text{(A.9.6)}$$

The population fractions can then be computed as:

$$Y_D = \frac{2[D]}{2[D] + 2[I_D] + [U]} = \frac{2K_{2U,I_D} K_{I_D,D}[U]^2}{2K_{2U,I_D} K_{I_D,D}[U]^2 + 2K_{2U,I_D}[U]^2 + [U]} \quad \text{(A.9.7)}$$

A10. Code for simulation and fitting of the analysis of holoMb unfolding by a six-state mechanism with heme dimerization

#Code should be in a text file and run in Gnuplot

#The data should be assembled into an excel sheet with the data gathered into a consecutive column. Column 1 should denote a reference to the protein concentration and detection method. For example, I used 2, 3, 10, 11, 100, 101 to denote the different data sets. Column 2 should be the [GuHCl] and the respective data should be arranged sequentially in Column 3. In order to create a multiplot for the normalized signals and population fractions, each data set was copied and arranged side by side in Columns 4 to 9. Once the sheet was complete, the excel sheet was converted into a text file (Windows format, not MS-DOS) for Gnuplot to be able to read it.

```
#This terminal should be used for working on the fits. However, the output only displays the fits but is not
is not readable as a ps file.
#set term x11 enhanced font "times,28"
#set output 'WTfits.ps'

#This terminal and output allows for very nice and readable figures
set terminal postscript eps enhanced color
set output "WTfitsforpub.eps"

set title 'Equilibrium Unfolding of WT HoloMb'

# Which protein concentration?
#This equation allows for obtaining the protein concentration out of x
P(x)=(floor(x/100));

# [GuHCl] Concentration
#This equation allows for obtaining the [GuHCl] out of x.
gc(x)=x-100*P(x);

# Defining the dependence on [denaturant]
knhv(x)=knh*exp(mnh*gc(x)/RT);
kihv(x)=kih*exp(mih*gc(x)/RT);
kuhv(x)=kuh*exp(muh*gc(x)/RT);
kniv(x)=kni*exp(mni*gc(x)/RT);
kiuv(x)=kiu*exp(miu*gc(x)/RT);
kDv(x)=kD*exp(mD*gc(x)/RT);

# Defining the cubic function with the a b c d coefficients
a(x) = 2.*kDv(x)*(1 + knhv(x)*kniv(x)/kihv(x) + knhv(x)*kniv(x)*kiuv(x)/kuhv(x));
```

```
c(x) = knhv(x) + knhv(x)*kniv(x) + knhv(x)*kniv(x)*kiuv(x);
b(x) = 2.*kDv(x)*c(x) + (1+knhv(x)*kniv(x)/kihv(x) + knhv(x)*kniv(x)*kiuv(x)/kuhv(x));
d2(x) = (-2e-6)*c(x);
d10(x) = (-10e-6)*c(x);
d100(x) = (-100e-6)*c(x);

#Changing to Monic
A(x) = b(x)/a(x);
B(x) = c(x)/a(x);
C2(x) = d2(x)/a(x);
C10(x) = d10(x)/a(x);
C100(x) = d100(x)/a(x);

#Solving the cubic and finding the correct root
Pea(x) = B(x) - A(x)**2/3.;

Que2(x) = C2(x) + (2*A(x)**3 - 9*A(x)*B(x))/ 27.;
You2(x) = (-Que2(x)/2. + (Que2(x)**2 /4. + Pea(x)**3 /27.)**(1./2.))**(1./3.);
h2(x) = You2(x) - Pea(x)/(3.*You2(x)) - A(x)/3.
Que10(x) = C10(x) + (2*A(x)**3 - 9*A(x)*B(x))/ 27.;
You10(x) = (-Que10(x)/2. + (Que10(x)**2 /4. + Pea(x)**3 /27.)**(1./2.))**(1./3.);
h10(x) = You10(x) - Pea(x)/(3.*You10(x)) - A(x)/3.
Que100(x) = C100(x) + (2*A(x)**3 - 9*A(x)*B(x))/ 27.;
You100(x) = (-Que100(x)/2. + (Que100(x)**2 /4. + Pea(x)**3 /27.)**(1./2.))**(1./3.);
h100(x) = You100(x) - Pea(x)/(3.*You100(x)) - A(x)/3.

# Partition Function
Q2(x)=1+knhv(x)/h2(x)+knhv(x)*kniv(x)/h2(x)+knhv(x)*kniv(x)*kiuv(x)/h2(x)+knhv(x)*kniv(x)/kihv(x)
+knhv(x)*kniv(x)*kiuv(x)/kuhv(x);

Q10(x)=1+knhv(x)/h10(x)+knhv(x)*kniv(x)/h10(x)+knhv(x)*kniv(x)*kiuv(x)/h10(x)+knhv(x)*kniv(x)/kih
v(x)+knhv(x)*kniv(x)*kiuv(x)/kuhv(x);

Q100(x)=1+knhv(x)/h100(x)+knhv(x)*kniv(x)/h100(x)+knhv(x)*kniv(x)*kiuv(x)/h100(x)+knhv(x)*kniv(x
)/kihv(x)+knhv(x)*kniv(x)*kiuv(x)/kuhv(x);

# Defining the population fractions
YNH2(x)=1/Q2(x);
YIH2(x)=(knhv(x)*kniv(x)/kihv(x))/Q2(x);
YUH2(x)=(knhv(x)*kniv(x)*kiuv(x)/kuhv(x))/Q2(x);
YN2(x)=(knhv(x)/h2(x))/Q2(x);
YI2(x)=(knhv(x)*kniv(x)/h2(x))/Q2(x);
YU2(x)=(knhv(x)*kniv(x)*kiuv(x)/h2(x))/Q2(x);
YH2(x)=h2(x)/(2e-6);
YD2(x)=2*kDv(x)*h2(x)*h2(x)/(2e-6);

YNH10(x)=1/Q10(x);
YIH10(x)=(knhv(x)*kniv(x)/kihv(x))/Q10(x);
YUH10(x)=(knhv(x)*kniv(x)*kiuv(x)/kuhv(x))/Q10(x);
YN10(x)=(knhv(x)/h10(x))/Q10(x);
YI10(x)=(knhv(x)*kniv(x)/h10(x))/Q10(x);
YU10(x)=(knhv(x)*kniv(x)*kiuv(x)/h10(x))/Q10(x);
YH10(x)=h10(x)/(10e-6);
YD10(x)=2*kDv(x)*h10(x)*h10(x)/(10e-6);

YNH100(x)=1/Q100(x);
```

```
YIH100(x)=(knhv(x)*kniv(x)/kihv(x))/Q100(x);
YUH100(x)=(knhv(x)*kniv(x)*kiuv(x)/kuhv(x))/Q100(x);
YN100(x)=(knhv(x)/h100(x))/Q100(x);
YI100(x)=(knhv(x)*kniv(x)/h100(x))/Q100(x);
YU100(x)=(knhv(x)*kniv(x)*kiuv(x)/h100(x))/Q100(x);
YH100(x)=h100(x)/(100e-6);
YD100(x)=2*kDv(x)*h100(x)*h100(x)/(100e-6);

#Input of temperature
TC=20; #Temperature in C
T=TC+273.15; #Temperature in Kelvin
R=0.00831447; # Ideal gas constant in kJ.mol-1.M-1 K-1
RT=R*T;

# Setting the observables
SCDNH=0.9933;
SCDN=0.6;
SCDI=0.25;
SCDIH=0.05715;
SCDU=0.0257;
SCDUH=0.01;

SABSNH=1.01152;
SABSIH=0.18810;
SABSUH=0.02;
SABSH=0.152169;
SABSD=0.451;

# Defining the constants in buffer
#Note that ".**" is the same as "e" for expressiong exponentials
knh=7.37e-14; #
kih=1.18e-11; #
kuh=1e-6; # (M)
kni=1.6*10.**-2;
kiu=1.1*10.**-2;
kD=3.85e6; #

# Defining the m values
mnh=15.8171; #kJ.mol-1.M-1
mih=13.4951; #kJ.mol-1.M-1
muh=8.0; #kJ.mol-1.M-1
mni=9.850; #kJ.mol-1.M-1
miu=5.680; #kJ.mol-1.M-1
mD=-4.8; #kJ.mol-1.M-1

# Setting up the observables
ObsS2(x)=YNH2(x)*SABSNH+YIH2(x)*SABSIH+YUH2(x)*SABSUH+YH2(x)*SABSH+YD2(x)*SAB
SD;
ObsS10(x)=YNH10(x)*SABSNH+YIH10(x)*SABSIH+YUH10(x)*SABSUH+YH10(x)*SABSH+YD10(
x)*SABSD;
ObsS100(x)=YNH100(x)*SABSNH+YIH100(x)*SABSIH+YUH100(x)*SABSUH+YH100(x)*SABSH+
YD100(x)*SABSD;
ObsCD2(x)=YNH2(x)*SCDNH+YIH2(x)*SCDIH+YUH2(x)*SCDUH+YN2(x)*SCDN+YI2(x)*SCDI+Y
U2(x)*SCDU;
ObsCD10(x)=YNH10(x)*SCDNH+YIH10(x)*SCDIH+YUH10(x)*SCDUH+YN10(x)*SCDN+YI10(x)*S
CDI+YU10(x)*SCDU;
```

```
ObsCD100(x)=YNH100(x)*SCDNH+YIH100(x)*SCDIH+YUH100(x)*SCDUH+YN100(x)*SCDN+YI1
00(x)*SCDI+YU100(x)*SCDU;

print ""
print "******"
print "Overall fitting showing dependence on protein concentration."
print "******"
print ""
print "******"
print "KNH= ", knh, " / ", "KIH= ", kih, " / ", "KUH= ", kuh, " / ", "KH2= ", kD
print "mNH= ", mnh, " / ", "mIH= ", mih, " / ", "mUH= ", muh, " / ", "mH2= ", mD
print "CDNH= ", SCDNH, " / ", "CDN= ", SCDN, " / ", "CDIH= ", SCDIH, " / ", "CDI= ", SCDI, " /
", "CDUH= ", SCDUH, " / ", "CDU= ", SCDU
print "AbsNH= ", SABSNH, " / ", "AbsIH= ", SABSIH, " / ", "AbsUH= ", SABSUH, " / ", "AbsH= ",
SABSH, " / ", "AbsH2= ", SABSD
print "******"

Observ(x) = \
(P(x)==2)?(ObsS2(x)):(P(x)==3)?(ObsCD2(x)):\
(P(x)==10)?(ObsS10(x)):(P(x)==11)?(ObsCD10(x)):\
(P(x)==100)?(ObsS100(x)):(ObsCD100(x))

#ObsS(x) = (P(x)==2)?(ObsS2(x)):(P(x)==10)?(ObsS10(x)):(ObsS100(x));
#ObsCD(x) = (P(x)==2)?(ObsCD2(x)):(P(x)==10)?(ObsCD10(x)):(ObsCD100(x));
#ObsF(x) = (P(x)==2)?(ObsF2(x)):(P(x)==10)?(ObsF10(x)):(ObsF100(x));

#Overall fitting of data
#FIT_MAXITER = 100
#FIT_LIMIT = 1e-9
#fit Observ(x) "WTdat2.txt" using (100*$1+$2):3 via SABSIH,SCDIH

#Plotting the fitting results with data
set multiplot
reset
set samples 500
set xrange [0:]
set yrange [-0.05:1.2]
set xlabel '[GuHCl] (M)'
set ylabel 'Normalized Signals'
set size 0.48,0.5
set origin 0.01,0.51
set title "Normalized Signals"
#set key right bottom
unset key
plot \
'WTdat2.txt' using 2:4 with points pt 7 lt 1, 'WTdat2.txt' using 2:5 with points pt 7 lt 3,\
'WTdat2.txt' using 2:6 with points pt 9 lt 1, 'WTdat2.txt' using 2:7 with points pt 9 lt 3,\
'WTdat2.txt' using 2:8 with points pt 5 lt 1, 'WTdat2.txt' using 2:9 with points pt 5 lt 3,\
 ObsS2(x) with lines lt 4 lc 1 lw 4, ObsS10(x) with lines lt 2 lc 1 lw 4, ObsS100(x) with lines lt -1 lc 1 lw
2,\
 ObsCD2(x) with lines lt 4 lc 3 lw 4, ObsCD10(x) with lines lt 2 lc 3 lw 4, ObsCD100(x) with lines lt -1 lc
3 lw 2

#plotting NH, IH, and UH
set xrange [0:4]
set yrange [-0.05:1.05]
```

```
set xlabel '[GuHCl] (M)'

set ylabel 'Population Fractions'
set origin 0.01,0.01
set title "Population fractions"
plot \
YNH2(x) with lines lt 4 lc 7 lw 4, YNH10(x) with lines lt 2 lc 7 lw 4, YNH100(x) with lines lt -1 lc 7 lw 2,\
YIH2(x) with lines lt 4 lc 8 lw 4, YIH10(x) with lines lt 2 lc 8 lw 4, YIH100(x) with lines lt -1 lc 8 lw 2,\
YUH2(x) with lines lt 4 lc 3 lw 4, YUH10(x) with lines lt 2 lc 3 lw 4, YUH100(x) with lines lt -1 lc 3 lw 2

#plotting N, I, and U
set xrange [0:4]
set yrange [-0.05:1.05]
set xlabel '[GuHCl] (M)'

set ylabel 'Population Fractions'
set origin 0.51,0.51
set title "Population fractions"
plot \
YN2(x) with lines lt 4 lc 9 lw 4, YN10(x) with lines lt 2 lc 9 lw 4, YN100(x) with lines lt -1 lc 9 lw 2,\
YI2(x) with lines lt 4 lc 2 lw 4, YI10(x) with lines lt 2 lc 2 lw 4, YI100(x) with lines lt -1 lc 2 lw 2,\
YU2(x) with lines lt 4 lc 1 lw 4, YU10(x) with lines lt 2 lc 1 lw 4, YU100(x) with lines lt -1 lc 1 lw 2

#plotting H and D
set xrange [0:4]
set yrange [-0.05:1.05]
set xlabel '[GuHCl] (M)'

set ylabel 'Population Fractions'
set origin 0.51,0.01
set title "Population fractions"
plot \
YD2(x) with lines lt 4 lc 7 lw 4, YD10(x) with lines lt 2 lc 7 lw 4, YD100(x) with lines lt -1 lc 7 lw 2,\
YH2(x) with lines lt 4 lc 9 lw 4, YH10(x) with lines lt 2 lc 9 lw 4, YH100(x) with lines lt -1 lc 9 lw 2

unset multiplot
reset
#to output fitting params to a text file
set print 'fitparams.txt'
print\
SABSNH,SABSIH,SABSH,SABSD,knh,kih,kuh,mnh,mih,muh,kD,mD,SCDNH,SCDN,SCDIH,SCDI,SC
DUH,SCDU,
unset print
```